

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 14:32:03 ; Search time 67 Seconds  
(without alignments)  
443.506 Million cell updates/sec

Title: US-09-817-199a-2

Perfect score: 1150

Sequence: 1 MTPTGAVATRDGEAPERSP.....FQIRDYVESQKRSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	100.0	232	AAU41696	Human polypeptide
2	1144	99.5	226	AAU17119	Novel signal trans
3	1133	98.5	222	AAU17541	Novel signal trans
4	977	85.0	191	AAG67156	Amino acid sequenc
5	977	85.0	191	AAU39910	Human polypeptide
6	977	85.0	191	ABU76426	Human Ras protein
7	969	84.3	191	AAU52190	Human degranulatio
8	946	82.3	191	AAU52189	Mouse degranulatio
9	940	81.7	191	AAU52188	Degradation regu
10	926	80.5	191	AAU30132	An exocytotic prot

11	770	67.0	158	22	ABG17244	Novel human diagno
12	697	60.6	666	22	ABU71578	Drosophila melanog
13	666	57.9	190	19	AAU80747	Human RAB protein,
14	535	46.5	139	22	AAU17123	Novel signal trans
15	514.5	44.7	221	23	ABU41333	Human ovarian anti
16	513.5	44.7	213	22	ABU1916	Human Rab8 homolog
17	509	44.3	218	21	AAU19220	Arabidopsis thalia
18	508.5	44.2	204	22	ABU70670	Drosophila melanog
19	507	44.1	253	21	AAU39944	Arabidopsis thalia
20	507	44.1	254	21	AAU80687	Arabidopsis thalia
21	507	44.1	258	21	AAU30498	Arabidopsis thalia
22	505.5	44.0	215	21	AAU35215	Zea mays protein f
23	502	43.7	216	21	AAU80806	Arabidopsis thalia
24	501	43.6	203	21	AAU30499	Arabidopsis thalia
25	501	43.6	207	22	AAU67154	Amino acid sequenc
26	501	43.6	207	22	AAU92628	Human protein sequ
27	500	43.5	216	21	AAU80688	Arabidopsis thalia
28	500	43.5	216	21	AAU3945	Arabidopsis thalia
29	498	43.3	218	21	AAU56993	Human prostate can
30	498	43.3	218	22	ABU7266	Novel human diagno
31	497.5	43.3	202	21	AAU10858	Arabidopsis thalia
32	496.5	43.2	202	21	AAU07763	Arabidopsis thalia
33	494	43.0	200	21	AAU19165	Amino acid sequenc
34	494	43.0	200	21	AAU09979	Human Rab10 protei
35	494	43.0	200	22	AAU95340	Human protein sequ
36	488	42.4	224	21	AAU47826	Arabidopsis thalia
37	488	42.4	234	21	AAU47825	Arabidopsis thalia
38	487.5	42.4	201	21	AAU09982	Canine Rab10 prote
39	487.5	42.4	207	22	ABU1647	Drosophila melanog
40	487	42.3	209	22	ABU23365	Novel human diagno
41	483	42.0	218	21	AAU30710	Arabidopsis thalia
42	483	42.0	221	21	AAU30709	Arabidopsis thalia
43	483	42.0	246	21	AAU58196	Lung cancer associ
44	481.5	41.9	199	21	AAU09980	Human Rab10 protei
45	481.5	41.9	199	21	AAU09981	Human Rab10 protei

ALIGNMENTS

RESULT 1	AAU41696	AAU41696 standard; Protein: 232 AA.
ID	AAU41696	AAU41696 standard; Protein: 232 AA.
XX	AAU41696	
AC	AAU41696	
XX	AAU41696	
DT	22-OCT-2001	(first entry)
XX	Human polypeptide SEQ ID NO 6627.	
DE	Human polypeptide SEQ ID NO 6627.	
XX	Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	WO200153312-A1.	
PN	WO200153312-A1.	
XX	26-JUL-2001.	
PD	26-JUL-2001.	
XX	26-DEC-2000; 2000WO-US34263.	
PF	26-DEC-2000; 2000WO-US34263.	
XX	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	

XX (HYSE-) HYSEQ INC.  
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;  
XX WPI; 2001-442253/47.  
DR N-PSDB; AAI60852.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 2; SEQ ID NO 6627; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAAM2213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 232 AA;  
Query Match 100.0%; Score 1150; DB 22; Length 232;  
Best Local Similarity 100.0%; Pred. No. 9.6e-120;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60  
Db 10 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 69  
QY 61 ATVGIDFRKNVTVGVRVKLQIWDTAGQERFSVTHAYYRQAALLLYDITNKSSFDN 120  
Db 70 ATVGIDFRKNVTVGVRVKLQIWDTAGQERFSVTHAYYRQAALLLYDITNKSSFDN 129  
QY 121 IRAWLTIHEYAQRDVVIMLLGNKADMSSEVIRSEDGETFLAREYGVPFLETSAKTGMNV 180  
Db 130 IRAWLTIHEYAQRDVVIMLLGNKADMSSEVIRSEDGETFLAREYGVPFLETSAKTGMNV 189  
QY 181 ELAFLAIKELKYRAGHQADEPFSQIRDYVESQKRSSCCSFM 223  
Db 190 ELAFLAIKELKYRAGHQADEPFSQIRDYVESQKRSSCCSFM 232  
RESULT 2  
AAU17119  
ID AAU17119 standard; Protein; 226 AA.  
XX  
AC AAU17119;  
XX  
XX 07-NOV-2001 (first entry)  
DT  
XX  
DE Novel signal transduction pathway protein, Seq ID 684.  
XX  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS;  
KW acquired immune deficiency syndrome.  
XX Homo sapiens.  
XX WO200154733-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01312.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0203515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0218880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0228279.  
PR 22-AUG-2000; 2000US-0228681.  
PR 22-AUG-2000; 2000US-0228688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.  
PR 26-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-463460/50.  
DR N-PSDB; AAS27036.  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders -  
XX  
PS Claim 1; SEQ ID No 684; 880pp; English.  
XX The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative  
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
CC Addison's disease), reproductive system disorders, gastrointestinal  
CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
CC as stimulators of B-cell responsiveness to pathogens, activators of  
CC T-cells, to induce higher affinity antibodies, and as a means to induce  
CC tumour proliferation in pathologies e.g. acquired immune deficiency  
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
CC pathway protein, amino acid sequences of the invention.  
XX

Query Match 99.5%; Score 1144; DB 22; Length 226;

Best Local Similarity 99.6%; Pred. No. 4.3e-119; Mismatches 1; Indels 0; Gaps 0;  
Matches 222; Conservative 0;

QY 1 MTGTPCAVATRDGEAPERSPSPSYDLTGKVMVLGDTGVGKTCFLIQKDFGLSGTFI 60

Db 4 MTGTPCAVATRDGEAPERSPSPSYDLTGKVMVLGDTGVGKTCFLIQKDFGLSGTFI 63

QY 61 ATVGIDFRNKVTVVDGVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 120

Db 64 ATVGIDFSNKVTVVDGVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 123

QY 121 IRAWLITEIHEYAQRDVIIMLGNKADMSSEVRIRSEDTGLAREYGVPLETSAKTGMNV 180

Db 124 IRAWLITEIHEYAQRDVIIMLGNKADMSSEVRIRSEDTGLAREYGVPLETSAKTGMNV 183

QY 181 ELAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 223

Db 184 ELAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 226

RESULT 3

AAU17541

ID AAU17541 standard; Protein; 222 AA.

XX AAU17541;

XX 07-NOV-2001 (first entry)

XX Novel signal transduction pathway protein, Seq ID 1106.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
immune system disorder; rheumatoid arthritis; inflammatory condition;

KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX WO200154733-A1.  
XX 02-AUG-2001.  
PF 17-JAN-2001; 2001WO-US01312.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235836.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.



PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX WPI: 2001-465460/50.  
 DR N-PSDB; AAS27458.  
 XX  
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders -  
 XX  
 PS Claim 1; SEQ ID No 1106; 880pp; English.  
 XX  
 CC The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
 CC pathway protein, amino acid sequences of the invention.  
 XX  
 Query Match 98.5%; Score 1133; DB 22; Length 222;  
 Best Local Similarity 99.1%; Pred No. 7, 1e-118;  
 Matches 220; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 TCTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFSLGTF 61  
 Db 1 TCTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFSLGTF 60  
 QY 62 TVGIDFRNKVTVVDGVRVKLQIWDTAGQFRFSVTHAYYRDAQALLLLYDITNKSSFDNI 121  
 Db 61 TVGIDFNKVVTVVDGVRVKLQIWDTAGQFRFSVTHAYYRDAQALLLLYDITNKSSFDNI 120  
 QY 122 RAWLTEIHEYAQRDVVIMLGNKADMSERVIRSEDETGLAREYGVVPFLETSAKTGMNVE 181  
 Db 121 RAWLTEIHEYAQRDVVIMLGNKADMSERVIRSEDETGLAREYGVVPFLETSAKTGMNVE 180  
 QY 182 LAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 223  
 Db 181 LAFLAIKELKYRAGHQADEPSFQIRDYVYKSQKRSSCCSFM 222  
 RESULT 4  
 AAG67156  
 ID AAG67156 standard; Protein; 191 AA.  
 XX  
 AC AAG67156;  
 XX  
 DT 13-NOV-2001 (first entry)  
 XX

DE Amino acid sequence of human 32712 G-protein.  
 XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;  
 KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;  
 KW adult respiratory distress syndrome; Goodpasture's syndrome;  
 KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;  
 KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;  
 KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;  
 KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;  
 KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;  
 KW chronic bacterial meningoencephalitis; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; stroke; Huntington's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164887-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 27-FEB-2001; 2001WO-US06292.  
 XX  
 PR 29-FEB-2000; 2000US-0185606.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Meyers RA;  
 XX  
 XX WPI: 2001-550182/61.  
 DR N-PSDB; AAH75184.  
 XX  
 PT Novel human small G-protein polypeptides and polynucleotides for  
 PT treating lung disorders, liver disorders and brain disorders -  
 XX  
 PS Claim 8; Fig 26; 151pp; English.  
 XX  
 CC The present sequence represents a human G-protein. The specification  
 CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The  
 CC G-protein polypeptides and polynucleotides are useful as a target for  
 CC diagnosis and treatment of G-protein mediated or related disorders,  
 CC and for identifying agonists and antagonists for diagnosis and  
 CC treatment. They are useful for treating disorders of lung (e.g.  
 CC congenital anomalies, pulmonary congestion, oedema, adult respiratory  
 CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,  
 CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver  
 CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,  
 CC Wilson's disease, autoimmune hepatitis and hepatic failure), and  
 CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute  
 CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic  
 CC bacterial meningoencephalitis, multiple sclerosis, amyotrophic lateral  
 CC sclerosis, stroke and Huntington's disease).  
 XX  
 SQ Sequence 191 AA;

Query Match 85.0%; Score 977; DB 22; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;  
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 MLIGDTGVGKTCFLIQKDGAFSLGTFIATVGIDFRNKVTVVDGVRVKLQIWDTAGQFR 92  
 Db 1 MLIGDTGVGKTCFLIQKDGAFSLGTFIATVGIDFRNKVTVVDGVRVKLQIWDTAGQFR 60  
 QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLGNKADMSERV 152  
 Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLGNKADMSERV 120  
 QY 153 IRSEDETGLAREYGVVPFLETSAKTGMNVELAFIAIAKELKYRAGHQADEPSFQIRDYVES 212  
 Db 121 IRSEDETGLAREYGVVPFLETSAKTGMNVELAFIAIAKELKYRAGHQADEPSFQIRDYVES 180  
 QY 213 QKRSSCCSFM 223  
 Db 181 QKRSSCCSFM 191

RESULT 5  
AAM39910  
ID AAM39910 standard; Protein; 191 AA.  
XX AC AAM39910;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 3055.  
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US34263.  
XX PR 21-JAN-2000; 2000US-0488725.  
XX PR 25-APR-2000; 2000US-0552317.  
XX PR 09-JUL-2000; 2000US-0598042.  
XX PR 19-JUL-2000; 2000US-0620312.  
XX PR 03-AUG-2000; 2000US-0653450.  
XX PR 14-SEP-2000; 2000US-0662191.  
XX PR 19-OCT-2000; 2000US-0693036.  
XX PR 29-NOV-2000; 2000US-0727344.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX DR N-PSDB; AAI59066.  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX Example 4; SEQ ID NO 3055; 10078pp; English.  
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC assays for receptor activity, cancer diagnosis and therapy, drug screening,  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX SQ Sequence 191 AA;  
Query Match 85.0%; Score 977; DB 22; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.5e-100;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 33 MLLGDTGVGKTCFLIQKDAFLSGTFIATVGFIDFRNKVTVDGVVRVKLQINDTAGQERF 92

Db 1 MLLGDTGVGKTCFLIQKDAFLSGTFIATVGFIDFRNKVTVDGVVRVKLQINDTAGQERF 60  
QY 93 RSVTHAYYRDAQALLLLYDITNKSSPDNIRAWLTHEIHYAQRDVVIMLLGNKADMSSERV 152  
Db 61 RSVTHAYYRDAQALLLLYDITNKSSPDNIRAWLTHEIHYAQRDVVIMLLGNKADMSSERV 120  
QY 153 IRSEGETLAREYGVFPFLETSARTGMNVELAFIAIAKELKYRAGHQADEPSPFOIRDYVES 212  
Db 121 IRSEGETLAREYGVFPFLETSARTGMNVELAFIAIAKELKYRAGHQADEPSPFOIRDYVES 180  
QY 213 QKKRSSCCSFM 223  
Db 181 QKKRSSCCSFM 191  
RESULT 6  
ABE76426  
ID ABB76426 standard; Protein; 191 AA.  
XX AC ABB76426;  
XX DT 27-AUG-2002 (first entry)  
XX DE Human Ras protein 3 (RASP-3).  
XX KW Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;  
KW immunosuppressive; antiinflammatory; signal transduction.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Region 1..19 /note= "transforming protein P21 family signature"  
FT Region 4..11 /note= "ATP/GTP-binding site, motif A (P-loop)"  
FT Region 22..38 /note= "transforming protein P21 family signature"  
FT Region 40..62 /note= "transforming protein P21 family signature"  
FT Region 63..80 /note= "transforming protein P21 family signature"  
FT Region /note= "Ran family signature sequence"  
FT Region 102..115 /note= "transforming protein P21 family signature"  
FT Region 130..168 /note= "Ran family signature sequence"  
FT Region 137..159 /note= "transforming protein P21 family signature"  
XX PN US6391580-B1.  
XX PD 21-MAY-2002.  
XX PF 08-MAY-1998; 98US-0075454.  
XX PR 12-DEC-1996; 96US-0766551.  
XX PA (INCY-) INCYTE GENOMICS INC.  
XX PI Hillman JL, Tang YT, Lal P, Guegler KJ, Corley NC, Patterson C;  
PI Batra S, Baughn MR;  
XX WPI; 2002-498774/53.  
XX DR N-PSDB; ABN83696.  
XX PT New isolated polynucleotides encoding Ras proteins designated RASP-1  
XX and RASP-4, for diagnosing, preventing and treating disorders  
XX associated with cell proliferation, particularly cancer and immune  
XX disorders -  
XX Example; Column 47-48; 34pp; English.  
XX CC The present sequence is the protein sequence of novel human Ras

CC protein 3 (RASP-3). RASP-3 shares 72% sequence identity with  
CC Rab26. It includes regions that resemble the signature sequences  
CC of the GTP-binding Ras superfamily, and family signatures of Ran,  
CC a subfamily of Ras proteins which function in nucleocytoplasmic  
CC transport, RNA synthesis, processing and export, and cell cycle  
CC checkpoint control, as well as family signatures of transforming  
CC protein p21, a subfamily of GTP-binding Ras proteins which have  
CC been implicated in a number of tumours. Northern analysis showed  
CC expression of RASP-3 in haematopoietic and immunological cDNA  
CC libraries, all of which were associated with inflammation and the  
CC immune response. The invention provides 7 novel human Ras proteins  
CC (RASP-1 to -7) and polynucleotides, expression vectors, host cells,  
CC antibodies, agonists and antagonists. It also provides methods for  
CC diagnosing, treating or preventing disorders associated with RASP  
CC expression, especially cancer and immune disorders.  
XX  
SQ Sequence 191 AA;

Query Match 85.0%; Score 977; DB 23; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.5e-100;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 MLLGDTGVGKTCFLIQFDGAFSLGTFIATVIGIDFRNKVTVVDGVRVQLQIWDITAGQERF 92  
Db 1 MLLGDTGVGKTCFLIQFDGAFSLGTFIATVIGIDFRNKVTVVDGVRVQLQIWDITAGQERF 60  
Qy 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152  
Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 120  
Qy 153 IRSEDETLAREYGVFPFLETSKATGMNVELAF LAIAKELKYRAGHQADEPSFQIRDYVES 212  
Db 121 IRSEDETLAREYGVFPFLETSKATGMNVELAF LAIAKELKYRAGHQADEPSFQIRDYVES 180  
Qy 213 QKKRSSCCSFM 223  
Db 181 QKKRSSCCSFM 191

RESULT 7  
AAM52190  
ID AAM52190 standard; Protein; 191 AA.  
XX  
AC AAM52190;  
XX  
DT 07-FEB-2002 (first entry)  
XX  
DE Human degranulation regulator SEQ ID NO 3.  
XX  
KW Degranulation; mast cell; human; mouse; antiallergic.  
OS Homo sapiens.  
XX  
FN WO200179478-A1.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-JP03268.  
XX  
PR 19-APR-2000; 2000JP-0118408.  
XX  
PA (DAIN ) DAINIPPON PHARM CO LTD.  
XX  
PI Yamada T, Ido M;  
XX  
DR WPI; 2002-041335/05.  
XX  
DR N-PSDB; ABA02774.  
XX  
PT Mast cell degranulation controller for treatment of allergies  
XX  
PS Claim 11; Page 64-65; 85pp; Japanese.  
XX  
CC The invention relates to a protein for regulating degranulation of mast

CC cells (degranulation regulators) and the encoding polynucleotides, with  
CC antiallergic activity, used in the treatment of allergies associated  
CC with degranulation of mast cells.  
XX  
SQ Sequence 191 AA;

Query Match 84.3%; Score 969; DB 23; Length 191;  
Best Local Similarity 99.0%; Pred. No. 1.1e-99;  
Matches 189; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 33 MLLGDTGVGKTCFLIQFDGAFSLGTFIATVIGIDFRNKVTVVDGVRVQLQIWDITAGQERF 92  
Db 1 MLLGDSGVGKTCFLIQFDGAFSLGTFIATVIGIDFRNKVTVVDGVRVQLQIWDITAGQERF 60  
Qy 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152  
Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 120  
Qy 153 IRSEDETLAREYGVFPFLETSKATGMNVELAF LAIAKELKYRAGHQADEPSFQIRDYVES 212  
Db 121 IRSEDETLAREYGVFPFLETSKATGMNVELAF LAIAKELKYRAGHQADEPSFQIRDYVES 180  
Qy 213 QKKRSSCCSFM 223  
Db 181 QKKRSSCCSFM 191

RESULT 8  
AAM52189  
ID AAM52189 standard; Protein; 191 AA.  
XX  
AC AAM52189;  
XX  
DT 07-FEB-2002 (first entry)  
XX  
DE Mouse degranulation regulator SEQ ID NO 2.  
XX  
KW Degranulation; mast cell; human; mouse; antiallergic.  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 83  
FT /note= "Encoded by CAG"  
XX  
PN WO200179478-A1.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-JP03268.  
XX  
PR 19-APR-2000; 2000JP-0118408.  
XX  
PA (DAIN ) DAINIPPON PHARM CO LTD.  
XX  
PI Yamada T, Ido M;  
XX  
DR WPI; 2002-041335/05.  
XX  
DR N-PSDB; ABA02773.  
XX  
PT Mast cell degranulation controller for treatment of allergies  
XX  
PS Claim 10; Page 63-64; 85pp; Japanese.  
XX  
CC The invention relates to a protein for regulating degranulation of mast  
CC cells (degranulation regulators) and the encoding polynucleotides, with  
CC antiallergic activity, used in the treatment of allergies associated  
CC with degranulation of mast cells.  
XX  
SQ Sequence 191 AA;

Query Match 82.3%; Score 946; DB 23; Length 191;  
Best Local Similarity 96.3%; Pred. No. 4.2e-97;

```
Matches 184; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVDSVRVKLIQWDTAGQERF 92
DB 1 MLLGDSGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVDSVRVKLIQWDTAGQERF 60
QY 93 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 152
DB 61 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 120
QY 153 IRSEDEGTILAREGVGFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFOIRDYVES 212
DB 121 IRSEDEGTILAREGVGFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFOIRDYVES 180
QY 213 QKKRSSCCSFV 223
DB 181 QKKRSSCCSFV 191

RESULT 9
AAM521188
ID AAM521188 standard; protein; 191 AA.
XX
AC AAM521188;
XX
DT 07-FEB-2002 (first entry)
XX
DE Degranulation regulator SEQ ID NO 1.
XX
KW Degranulation; mast cell; human; mouse; antiallergic.
XX
OS Homo sapiens.
XX
Mus musculus.
XX
FH Key Location/Qualifiers
FT Misc-difference 45 /label= Ala, Val
FT Misc-difference 115 /label= Val, Met
FT Misc-difference 138 /label= Met, Leu
FT Misc-difference 165 /label= Arg, His
FT Misc-difference 167 /label= Pro, Ala
XX
XX WO200179478-A1.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-JP03268.
XX
XX 19-APR-2000; 2000JP-0118408.
XX
XX (DAIN ) DAINIPPON PHARM CO LTD.
XX
XX Yamada T, Ido M;
XX
XX WPI; 2002-041335/05.
XX
XX Mast cell degranulation controller for treatment of allergies -
XX
PS Claim 1; Page 63; 85pp; Japanese.
XX
XX The invention relates to a protein for regulating degranulation of mast
XX cells (degranulation regulators) and the encoding polynucleotides, with
XX antiallergic activity, used in the treatment of allergies associated
XX with degranulation of mast cells.
XX
XX Sequence 191 AA:
SQ
Query Match 81.7%; Score 940; DB 23; Length 191;
Best Local Similarity 96.3%; Pred. No. 2e-96;
```

```
Matches 184; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVDSVRVKLIQWDTAGQERF 92
DB 1 MLLGDSGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVDSVRVKLIQWDTAGQERF 60
QY 93 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 152
DB 61 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 120
QY 153 IRSEDEGTILAREGVGFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFOIRDYVES 212
DB 121 IRSEDEGTILAREGVGFLETSKATGMNVELAFIAIAKELKYRAGHQADEPSFOIRDYVES 180
QY 213 QKKRSSCCSFV 223
DB 181 QKKRSSCCSFV 191

RESULT 10
AAY30132
ID AAY30132 standard; Protein; 191 AA.
XX
AC AAY30132;
XX
DT 27-OCT-1999 (first entry)
XX
DE An exocytotic protein designated Exo2.
XX
KW Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
XX inflammation.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /note= "encoded by GAN"
FT Misc-difference 27 /note= "encoded by TTN"
XX
XX WO9942586-A2.
XX
XX 26-AUG-1999.
XX
XX 23-FEB-1999; 99WO-US03909.
XX
XX 26-MAY-1998; 98US-0086650.
XX
XX 23-FEB-1998; 98US-0075534.
XX
XX (RIGE-) RIGEL PHARM INC.
XX
XX Anderson D, Fisher J, Huang B, Lorens J, Luo Y;
XX Shen M;
XX
XX WPI; 1999-518605/43.
XX
XX N-PSDB; AAX86720.
XX
XX New exocytotic proteins useful for diagnosis and treatment of
XX exocytosis-mediated conditions and in drug screening
XX
XX Claim 12; Fig 7; 53pp; English.
XX
XX The present sequence represents an exocytotic protein designated Exo2.
XX The protein has an inhibitory effect on exocytosis (i.e. the fusion of
XX secretory vesicles with the cellular plasma membrane) and can be
XX administered therapeutically to treat or prevent exocytosis-mediated
XX disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
XX undesirable release of compounds via exocytosis e.g. inflammation
XX mediated by the release compounds such as histamine. Exo2 is useful
XX diagnostically and to produce antibodies useful to purify the proteins
XX and therapeutically to reduce or eliminate the biological activity of
XX the protein. Exo2 nucleic acids can be used therapeutically to increase
XX Exo2 activity in cells by known gene therapy techniques. They can also
```

CC be used to produce probes or primers to isolate Exo2 proteins from other  
CC organisms, especially humans. The nucleic acids, host cells and proteins  
CC are useful in screening assays to identify binding agents, especially  
CC drug screening assays to identify agonists and antagonists useful  
CC therapeutically to enhance or reduce Exo2 activity.  
XX  
SQ Sequence 191 AA;  
  
Query Match 80.5%; Score 926; DB 20; Length 191;  
Best Local Similarity 94.8%; Pred. No. 7.2e-95;  
Matches 181; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
  
QY 33 MLGGDTGVTCTFLQKDGAFLSGTFIATVGIDFRNKVTVVGVVVKLQIWDTAGQERF 92  
Db 1 MLGGSGVGTCTFLQKDGAFLSGTFIATVGIDFRNKVTVVGVVVKLQIWDTAGQERF 60  
  
QY 93 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLITEIHEYAQRDVVIMLGNKADMSERV 152  
Db 61 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLITEIHEYAQRDVVIMLGNKADMSERV 120  
  
QY 153 IRSEGETLAREYGVVFFLETSAGTKGMNVELAFALAIKELKYRAGHQADPEFSQIRDYVES 212  
Db 121 IRSEGETLAREYGVVFFLETSAGTKGMNVELAFALAIKELKYRAGHQADPEFSQIRDYVES 180  
  
QY 213 QKKRSCCSFM 223  
Db 181 QKKRSCCSFV 191  
  
RESULT 11  
ABG17244  
ID ABG17244 standard; Protein; 158 AA.  
XX  
AC ABG17244;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #17235.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS81431.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 20; SEQ ID NO 47603; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 158 AA;  
  
Query Match 67.0%; Score 770; DB 22; Length 158;  
Best Local Similarity 99.3%; Pred. No. 1.4e-77;  
Matches 148; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MTGTPCAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60  
Db 10 MTGTPCAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 69  
  
QY 61 ATVGIDFRNKVTVVGVVVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFDN 120  
Db 70 ATVGIDFRNKVTVVGVVVKLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFDN 129  
  
QY 121 IRAWLITEIHEYAQRDVVIMLGNKADMS 149  
Db 130 IRAWLITEIHEYAQRDVVIMLGNKADMS 158  
  
RESULT 12  
ABB71578  
ID ABB71578 standard; Protein; 666 AA.  
XX  
AC ABB71578;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 41526.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL15681.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 41526; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 666 AA;

Query Match 60.6%; Score 697; DB 22; Length 666;  
Best Local Similarity 59.8%; Pred. No. 1.6e-68;  
Matches 137; Conservative 33; Mismatches 49; Indels 10; Gaps 3;

QY 3 GTPCAVAT---RDGEA-----PERSPPCSPSYDLTGKVMILGDTGVGKTCFLIQKDG 53  
DB 433 GSEGASATLCKNAGRALIRMISSKAPKEEDEFKIMLGDSGVGKTSLLIRFERDGR 492  
QY 54 FLSGTATVGDIFRNKVVTDGVVRVKLQIWDTAGOERFSVTHAYYRQAALLLYDIT 113  
DB 493 YVPSYFLSTVGDIFRNKVVVDGTRVKLQIWDTAGOERFSVTHAYYRDAHALLLYDVT 552  
QY 114 NKSSFNIRAWLTEIHEYAQRDVVIMLLGNKADMS-SERVIRSEGETTLAREYGVPPLET 172  
DB 553 NKTTYDNIRAWLGEIYEAQEDVVIVLIGNKADCSGSRQVKREDGERLGRHNVPFMET 612  
QY 173 SAKTGMNVELAFIAIAKELKYRAGHQADEPSPFOIRDYVSEQKRSSCCS 221  
DB 613 SAKTGLNVELSFYAVARQLKSRGHEGDDGKFNVDVRDNTKARSVCA 661

RESULT 13  
AAW80747  
ID AAW80747 standard; Protein; 190 AA.

AC AAW80747;  
XX  
XX  
DT 17-DEC-1998 (first entry)  
XX  
DE Human RAB protein, SRAB.  
XX  
KW Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;  
KW Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;  
KW viral infection; therapy; autoimmune sialosis; cystic fibrosis;  
KW diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;  
KW hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;  
KW polycystic renal disease.

OS Homo sapiens.  
XX  
XX WO9842839-A1.  
XX  
XX 01-OCT-1998.  
XX  
XX 25-MAR-1998; 98WO-US05871.  
XX  
XX 26-MAR-1997; 97US-0824873.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Guegler KJ, Hillman JL;  
XX  
XX WPI; 1998-542282/46.  
XX  
XX N-PSDB; AAV65197.

XX New human Rab protein, SRAB - useful for treating disorders  
PT associated with SRAB expression, including vesicle trafficking,  
PT viral infection, and cancer  
XX  
XX Claim 1; Fig 1; 31pp; English.

XX  
CC This sequence is the human RAB protein, SRAB, of the invention. The  
CC SRAB protein can be used in compositions useful in the diagnosis,  
CC prevention, or treatment of disorders associated with vesicle trafficking  
CC (including epilepsy, Huntington's disease, Parkinson's disease and  
CC schizophrenia), cancer, or viral infection. Host cells containing the DNA  
CC sequence can be used to produce SRAB recombinantly. The viral infections  
CC that can be treated using SRAB include those caused by retroviruses  
CC particularly HIV and HTLV, hepadnaviruses, particularly hepatitis C,  
CC hantaviruses, herpesviruses, and arboviruses. The cancers that can be  
CC treated include adenocarcinoma, leukaemia, lymphoma, melanoma and  
CC sarcoma, particularly cancers of the endocrine, gastrointestinal and  
CC nervous systems and cancers of the adrenal gland, brain, breast, colon,  
CC oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,  
CC prostate, salivary gland, stomach, thyroid, and uterus. The protein can  
CC also be used to treat conditions such as autoimmune sialosis,  
CC choroideraemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,  
CC hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme  
CC deficiency, pancreatitis, and polycystic renal disease.

XX Sequence 190 AA;

Query Match 57.9%; Score 666; DB 19; Length 190;  
Best Local Similarity 67.0%; Pred. No. 7.4e-66;  
Matches 126; Conservative 25; Mismatches 37; Indels 0; Gaps 0;

QY 33 MLIGDGTGVGKTCFLIQKDGAFLSGTFIATVGDIFRNKVVTVGVRVKLQIWDTAGOERF 92  
DB 1 MLVGDGSGVKTCLLGAIGWCFFGGDLHLHRSIDFRNKVLDDGVGVKVLQMDWTAGOERF 60  
QY 93 RSVTHAYYRDAHALLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNKADMSERY 152  
DB 61 RSVTHAYYRDAHALLLYDITNKASFQNIQAWLTEIHEYAQRDVVIMLLGNKADMSERY 120  
QY 153 IRSEGETTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRAGHQADEPSPFOIRDYVES 212  
DB 121 VKREDGEKLAKEYGLPFPMTSAKTGLNVLDAFTAIKELQKRSKMAPSEPRFRLHDYVKR 180  
QY 213 QKKRSSCC 220  
DB 181 EGRGASCC 188

RESULT 14  
AAU17123  
ID AAU17123 standard; Protein; 139 AA.  
XX  
XX AAU17123;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
XX Novel signal transduction pathway protein, Seq ID 688.  
XX  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell-anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
KW acquired immune deficiency syndrome.

XX Homo sapiens.  
OS  
XX WO200154733-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01312.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR

PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 11-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225577.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226686.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0227009.  
 PR 01-SEP-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-465460/50.  
 DR N-PSDB; AAS27040.  
 XX  
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cancers, immune  
 PT disorders and neuronal disorders





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 17, 2002, 16:41:03 ; Search time 2651 Seconds  
(without alignments)  
2448.105 Million cell updates/sec

Title: US-09-817-199A-2  
Perfect score: 1150  
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKRSSCCSFM 223

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet\_p2n model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09817199/runat\_13112002\_134902\_5165/app.query.fasta\_1.391  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817199 @CGN\_1.1.1616 @runat\_13112002\_134902\_5165 -NCPU=6 -ICPU=3  
-NO\_XLPYX -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: gb\_wa.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1150	100.0	1700	9	BC016615 Homo sapi
2	1150	100.0	1977	9	AK098068 Homo sapi
3	1145	99.6	1116	6	AX236082 Sequence
4	1081	94.0	690	10	AF233582 Mus muscu
5	977	85.0	576	6	AX236084 Sequence
6	780.5	67.9	1513	9	BC007681 Homo sapi
7	757	65.8	1320	9	AB027137 Homo sapi
8	754	65.6	1098	10	RN018771 U18771 Rattus norv
9	753	65.5	573	9	AF498952 Homo sapi
10	731	63.6	1340	6	AK062279 Sequence
11	710.5	61.8	3105	9	AK054846 Homo sapi
12	649.5	56.5	2536	3	AY061826 Drosophil
13	521.5	45.3	1337	10	BC019990 Mus muscu
14	518.5	45.1	1074	8	AB024994 Cicer ari
15	517.5	45.0	2145	3	DDISAS1A M34456 D.discoidu
16	517.5	45.0	10034	2	AC115606 Dictyoste
17	516	44.9	638	10	S53270 MEL-RAS-rel
18	516	44.9	765	5	DYGORAZ2 M38391 Discopyge o
19	516	44.9	911	8	AY084345 Arabidops
20	516	44.9	2048	9	BC002977 Homo sapi
21	516	44.9	2818	9	AK025165 Homo sapi
22	513.5	44.7	624	9	AF498943 Homo sapi
23	513.5	44.7	660	9	HSRAB8 X56741 H.sapiens m
24	513.5	44.7	740	10	RN053475 U53475 Rattus norv
25	513.5	44.7	760	4	CFRAB8 X56385 Canine rab8
26	513.5	44.7	782	8	PEAGTPBP09 D12548 Pea mRNA fo
27	513.5	44.7	1265	9	AB038995 Homo sapi
28	512.5	44.6	1080	8	BVRAB1 Z49152 B.vulgaris
29	512	44.5	1980	9	S53268 Homo sapien
30	511.5	44.5	2497	6	AX285074 Sequence
31	511.5	44.5	2497	6	AX285080 Sequence
32	511.5	44.5	2497	6	AX285089 Sequence
33	511.5	44.5	2497	6	AX285091 Sequence
34	510	44.3	919	8	DCJ001367 AJ001367 Daucus ca
35	509.5	44.3	1082	8	LJRA8D J73947 L.japonicus
36	509	44.3	955	8	S66160 r1c1-eras-re
37	509	44.3	1163	8	LJRA8A J73944 L.japonicus
38	508.5	44.2	1613	3	AB006189 Drosophil
39	508.5	44.2	2349	3	AY060425 Drosophil
40	507.5	44.1	979	8	LJRA1C J73932 L.japonicus
41	507	44.1	681	8	AY056351 Arabidops
42	507	44.1	821	8	AY052333 Arabidops
43	507	44.1	986	8	AB079023 Nicotiana
44	507	44.1	1025	8	AY086756 Arabidops
45	507	44.1	1033	8	AB079021 Nicotiana

ALIGNMENTS

```

BC016615
LOCUS      BC016615      1700 bp      mRNA      linear      PRI 05-NOV-2001
DEFINITION Homo sapiens, similar to RAB37, member of RAS oncogene family,
            clone MGC:21391 IMAGE:4520191, mRNA, complete cds.
ACCESSION  BC016615
VERSION     BC016615.1 GI:16741620
KEYWORDS   MGC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1700)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (31-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-re@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E.-E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalona@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
            Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E.-E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 28 Row: J Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. .1700
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="MGC:21391 IMAGE:4520191"
   /tissue_type="Prostate, adenocarcinoma."
   /clone_lib="NIH_MGC-91"
   /lab_host="DH10B"
   /note="Vector: pCMV-SPORT6"
3. .674
   /codon_start=1
   /product="Similar to RAB37, member of RAS oncogene family"
   /protein_id="AAH16615.1"
   /db_xref="GI:16741621"
   /translation="MTGPGAVATRDGEAPERSPPSPSYDLTGKVMILGDTGVGKTC
FLIQKGFALSCFTATGVDFRNKVTVDGVKVLQIWDTAQCEFRSFWTHAYRD
AQALLLDYTNKSFNFIRNLWTEIHEYAQRDVVIMLGNKADMSSEVIRSDGET
LRAGYVFFLETSAKTMGNVELAFALAKELKIRAGHQADPEFSQIRDYVESQKRSS
CCSPM"

BASE COUNT 437 a 444 c 460 g 359 t
ORIGIN
Alignment Scores:
Pred. No.: 1.77e-108 Length: 1700
Score: 1150.00 Matches: 223
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-817-199A-2 (1-223) x BC016615 (1-1700)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
|||||

```

```

Db 3 ATGACGGGACGCCACGGCCCGCTGCCACCGGGATGGCGAGGCCGCCCGCTCCCG 62
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuGlyAspThrGlyVal 40
|||||
Db 63 CCTCGAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGCT 122
Qy 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
|||||
Db 123 GCGAAACATGTTTCTGATCAATTCAAAGACGGGGCTTCTGTCGCGAACCCTTCA 182
Qy 61 AlaThrValGlyLeuAspPheArgLysValValThrValAspGlyValArgValLys 80
|||||
Db 183 GCCACCGTCGGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTAGAGTGA 242
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 100
|||||
Db 243 CTCGAGATCTGGGACACCGCTGGGAGGAACGGTTCCGAAGCGTCAACCATGCTTAT 302
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
|||||
Db 303 AGAGATGCTCAGCGCTTGTCTCTGCTGATGACATCACCAACAATCTTCTTCGACA 362
Qy 121 IleArgAlaThrPheLeuThrGluLeuHisGluTyrAlaGlnArgAspValValMetLeu 140
|||||
Db 363 ATCAGGGCGCTGGCTCACTGAGATTGATGATGATGATGATGATGATGATGATG 422
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
|||||
Db 423 CTAGGCAACAAGCGGATATGACAGCGAAGAGATGATCCGTTCCGAGACCGAGAGAC 482
Qy 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
|||||
Db 483 TTGGCCAGGGAGTACGGTGTCTCCCTGCGAGACCGCGCAAGACTGCCATGAATGT 542
Qy 181 GluLeuAlaPheLeuAlaLeuLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
|||||
Db 543 GAGTAGCCTTTCTGCCATCGCCCAAGGAACATAACCGGGCGGGCATCAGCGGAT 602
Qy 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCys 220
|||||
Db 603 GAGCCAGCTTCCAGATCCGAGCTATGTAGATGCCAGAGAGCGGCTCCAGCTGCTGC 662
Qy 221 SerPheMet 223
|||||
Db 663 TCCTTCATG 671

RESULT 2
AK098068
LOCUS      AK098068      1977 bp      mRNA      linear      PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
            to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION  AK098068
VERSION     AK098068.1 GI:21758001
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens trachea cDNA to mRNA, clone_lib:TRACH2
            clone:TRACH2000540.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Ninomiya,K., Wagatsuna,M., Kanda,K., Kondo,H., Yokoi,T.,
            Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
            Kamiyama,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
            Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
            Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
            Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 1977)
            Isogai,T. and Yamamoto,J.

```

**TITLE** Direct Submission  
**JOURNAL** Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
**COMMENT** NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and FAB; annotation: HRI and RAB.

**FEATURES**  
 Location/Qualifiers

1..1977  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="TRACH2000540"  
 /tissue\_type="trachea"  
 /clone\_lib="TRACH2"  
 /note="cloning vector: pME18SFL3"

CDS  
 57..728  
 /note="unnamed protein product"

/codon\_start=1  
 /protein\_id="BAC05227.1"  
 /db\_xref="GI:21758002"

/translation="MTGTPGAVATRDGEAPSPSPSYDLTGKVMLLGDTGVGKTC  
 FLIOFKDAFLSGTFIATVGVDFRNKVVTVDPVGRVKLIQIDTAGOERFSVTHAYRD  
 AOALLLYDINKSFDNIRAWLTHIHEYAQRDVVIMLLGNKADMSRVSRSDEGT  
 LAREYGVPLETSKTKGNMVELAFALAKELKYRAGHQAEPFQIRDIVESQKRSS  
 CCSFM"

**BASE COUNT** 442 a 560 c 539 g 436 t

**Alignment Scores:**

**Pred. No.:** 2,16e-108 **Length:** 1977  
**Score:** 1150.00 **Matches:** 223  
**Percent Similarity:** 100.00% **Conservative:** 0  
**Best Local Similarity:** 100.00% **Mismatches:** 0  
**Query Match:** 100.00% **Indels:** 0  
**DB:** 9 **Gaps:** 0

US-09-817-199A-2 (1-223) x AK098068 (1-1977)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20

Db 57 ATGACGGGCACGCCAGGCCGCTGCCACCCGGATGGGAGGCCGCCGAGCGCTCCCGG 116

Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40

Db 117 CCCTGCAGTCCGAGCTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 176

Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60

Db 177 GCGAAACATGTTCTGTATCCCAATTCAAAGACGGGGCTTCTGTCGGAACCTTCATPA 236

Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80

Db 237 GCCACCGTCGGCATAGACTTCAGAACAGGTGGTGACTGTGATGGCGTGAGAGTGAG 296

Qy 81 LeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100

Db 297 CTGCAGATCTGGGACACCGCTGGCAGGAACGGTTCGGAAGCGTCAACCATGCTATTAC 356

Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120

Db 357 AGATGCTCAGCGCTTGTCTTGTGTATGACATCACCAACAAATCTTCTTTCGACAC 416

Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140

Db 417 ATCAGGGCTTGCTCACTGAGATTCATGATATGCCAGAGGACGCTGCTGATCATGCTG 476

Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160

Db 477 CTAGCAACAAGCGGATATGAGCAGCAAGAGAGTATCCGTCCGAAGACGAGACAC 536

Qy 161 LeuAlaAaGgLUtYrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180

Db 537 TTGGCCAGGGAGTAGGTGTTCCCTTCTTGGAGACAGCCCAAGACTGGCATGATGTG 596

Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200

Db 597 GAGTTAGCTTCTTGGCCATCCCAAGGAACAGAAATACCGGCGGCATCAGCGGAT 656

Qy 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerCysCys 220

Db 657 GAGCCAGCTCCAGATCCGAGACTATGATAGTCCCAAGAACGCTCCAGCTGCTGC 716

Qy 221 SerPheMet 223

Db 717 TCCTTCAAG 725

**RESULT 3**

**LOCUS** AX236082 1116 bp DNA linear PAT 26-SEP-2001

**DEFINITION** Sequence 13 from Patent WO0164887.

**ACCESSION** AX236082

**VERSION** AX236082.1 GI:15795889

**KEYWORDS**

**SOURCE** human;

**ORGANISM** Homo sapiens

**REFERENCE** 1 (bases 1 to 1116)  
 AUTHORS Meyer, R.A.  
 TITLE Patent: WO 0164887-A 13 07-SEP-2001;  
 JOURNAL Millennium Pharmaceuticals, Inc. (US)

**FEATURES**  
 Location/Qualifiers  
 source 1..1116  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 124..699  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC88255.1"  
 /db\_xref="GI:15795890"  
 /translation="MLGDTGVGKTCFLQFKDGAFLSGTFTATVGIDFRNKVTVTDG  
 VRVKLIQIDTAGOERFSVTHAYRDAOALLLYDINKSFDNIRAWLTHIHEYAQR  
 DVVIMLLGNKADMSRVSRSDEGTLEAREYGVPLETSKTKGNMVELAFALAKEL  
 YRAGHQAEPFQIRDIVESQKRSSCCSFM"

**BASE COUNT** 250 a 313 c 322 g 221 t

**ORIGIN**

**Alignment Scores:**

**Pred. No.:** 3,35e-108 **Length:** 1116  
**Score:** 1145.00 **Matches:** 222  
**Percent Similarity:** 100.00% **Conservative:** 0  
**Best Local Similarity:** 100.00% **Mismatches:** 0  
**Query Match:** 99.57% **Indels:** 0  
**DB:** 6 **Gaps:** 0

US-09-817-199A-2 (1-223) x AX236082 (1-1116)

Qy 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21

Db 31 ACGGGCACGCCAGGCCGCTGCCACCCGGATGGGAGGCCGCCGAGCGCTCCCGCGCC 90

Qy 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41

Db 91 TGCAGTCCGAGCTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGCGCTGCGC 150

Qy 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61

Db 151 AAAACATGTTCTGTATCCCAATTCAAAGACGGGGCTTCTGTCGGAACCTTCATAGCC 210

Qy 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81

```
|||||
Db 211 ACCCTGGCATAGACTTTCAGAAACAGGTGGTGAAGCTGAGAGCTG 270
QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 271 CAGATCTGGACACCGCTGGCGAGAACGGTTCGGAAGCGTCACCCATGCTTTATTACAGA 330
QY 102 AspAlaGlnAlaLeuLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 331 GATGCTCAGCGCTTTCCTGCTGTATGACATCACCAACAATCTCTTTCGACACATC 390
QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeuLeu 141
Db 391 AGGCGCTGGCTCACTGAGATTCATGATGATGCCAGAGGACGTGGTGTGATCATCTGCTA 450
QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db 451 GGCACACAGCGGATATGACACAGCAAGAGTATCGTTCGGAAGCGGAGAGACCTTG 510
QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db 511 GCCAGGAGTACGTGTCTCTTCTGGACAGCGCCAAAGACTGGCATGAATGTGGAG 570
QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db 571 TTAGCCTTCTGGCCATCGCAAGGAAGTGAATACCGGGCGGCGCATCAGCGGATGAG 630
QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSer 221
Db 631 CCAGCTTCAGATCCGAGACTATGATAGATCCAGAGAGCGCTCCAGAGAGCGCTCC 690
QY 222 PheMet 223
Db 691 TTCATG 696

RESULT 4
AF233582 690 bp mRNA linear ROD 02-MAY-2000
LOCUS Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
DEFINITION
ACCESSION AF233582
VERSION AF233582.1 GI:7677421
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Masuda,E.S., Luo,Y., Young,C., Shen,M., Rossi,A.B., Huang,B.C.,
Yu,S., Bennett,M.K., Pavan,D.G. and Scheller,R.H.
TITLE Rab37 is a novel mast cell specific GTPase localized to secretory
granules
JOURNAL FEBS Lett. 470 (1), 61-64 (2000)
MEDLINE 20189834
PUBMED 10722846
REFERENCE
AUTHORS Luo,Y., Huang,B.C.B., Yu,S., Shen,M. and Masuda,E.S.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
Avenue, South San Francisco, CA 94080, USA
FEATURES
Source 1..690
/organism="Mus musculus"
/db_xref="taxon:10090"
1..690
/gene="Rab37"
1..672
/gene="Rab37"
/codon_start=1
/product="GTPase Rab37"
/protein_id="AAF67162.1"
/db_xref="GI:7677422"
/translation="WTGTPGATGAGAPSPFPSPNYDLTGKVMLLGDSGVKTC
FLIQFDAGFLSGTITATVIGIDFRNKVTVVDGARVKLQIWDTAGOFRFSVTHAYRD
```

```
ACQALLLYDITNQSSFDNIRAWLTIETHEYAQRDWWIMLLGNKADYSSESRVISEDGET
LAREYGVPPMETSAKTMNVELAFIAIAKELKVRAGQDPDFSQIRDVVEQKFRSS
CCSFV"
BASE COUNT 160 a 181 c 201 g 148 t
ORIGIN
Alignment Scores: 6.63e-102 Length: 690
Pred. No.: 1081.00 Matches: 209
Score: 96.41% Conservatives: 6
Percent Similarity: 93.72% Mismatches: 8
Best Local Similarity: 94.00% Indels: 0
Query Match: 10 Gaps: 0
DB:
US-09-817-199a-2 (1-223) x AF233582 (1-690)
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 1 ATGACTGGCACACCGAGAGCTGTACCGCTGGGGATGGCGAGGCCCTGTAGCGCTCCCG 60
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 61 CCCTTCAGCCCGACACGATCTACCGGCAAGGTATGCTCTTGGAGACTCGGGCGTC 120
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 121 GCGAAACCTGTTCTCTGATCAATTCAGAGACGGGCGCTTCCTGTCCGGAACCTTCATA 180
QY 61 AlaThrValGlyIleAspPheArgAsnLysValIleThrValAspGlyValArgValLys 80
Db 181 GCCACCGCTGGCATAGACTTCAGGAATAAGTGTGACAGTGGATGTGCGCAGGGTGAAG 240
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 241 CTCAGATCTGGGACACTCGCAGGACAGGAGCGCTTCGCGAGTGTGACCATGCTTATTAC 300
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 301 CGAGATGCTCAGCGCTTGTCTCTGTGTATGACATCACCAACCACTCTCTTTTGAACAAC 360
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140
Db 361 ATCAGGCGCTGGCTCACAGATTCATGATGATGCCAGAGGACGTGGTGTATGCTT 420
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 421 CTAGGCAACAACGCCGATGTAGCAGCAAGAGGTGATCCGTTCTGAGATGGAGAGACA 480
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db 481 CTGGCCAGGGAATATGCTGTTCTCTTCATGGAGACAGTCCCAAGACTGCGCATGAACGTG 540
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db 541 GAGTTGGCCTTCTGGCAATTGCCAAGGAACCTGAATACCGTCAGGGAGGACGCTGAT 600
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 601 GAGCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAAGAGGCGCTCCAGCTGCTGTC 660
QY 221 SerPheMet 223
Db 661 TCCTTTGTG 669
RESULT 5
AX236084
LOCUS AX236084
DEFINITION Sequence 15 from Patent WO0164887.
ACCESSION AX236084
VERSION AX236084.1 GI:15795891
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 576)  
AUTHORS Meyers R.A.  
TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins  
JOURNAL Patent: WO 0164887-A 15 07-SEP-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..576  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 141 a 148 c 167 g 120 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,48e-91 Length: 576  
Score: 977.00 Matches: 191  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 84.96% Indels: 0  
DB: 6 Gaps: 0  
US-09-817-199A-2 (1-223) x AX236084 (1-576)  
QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGly 52  
Db 1 ATGCTTCTGGGAGACAGCGCTCGCAAAACATGTTCTGATCCAAATTCAGAGCGG 60  
QY 53 AlaPheLeuSerGlyThrPheLeuAlaThrValGlyLeuAspPheArgAsnLysValVal 72  
Db 61 GCCTTCTGCGGAGACCTTATAGCACCGCTGGCATAGCTTCAGGAACAGGTGGTG 120  
QY 73 ThrValAspGlyValArgValLysLeuGlnLeuTyrPheThrAlaGlyGlnGluArgPhe 92  
Db 121 ACTGTGATGGCTGAGATGAGCTGCAGATCTGGGACACCGCTGGGAGGACCGTTC 180  
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112  
Db 181 CGAAGCGTCACCATGCTTATTACAGAGATGCTTCAGGCGCTTCTGCTGTATGATC 240  
QY 113 ThrAsnLysSerPheAspAsnIleArgAlaTyrPheThrGluLeuHisGluTyrAla 132  
Db 241 ACCAACAATCTCTTTCGACACATCAGGCGCTGGCTCAGATTCATGATGATGCC 300  
QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152  
Db 301 CAGAGGAGCGTGGTGATGCTGCTAGCAACAAGCGGATATGACGCGAAGAGTG 360  
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172  
Db 361 ATCCGTTCGAGACGAGAGACCTTGGCCAGGAGTACGGTGTCCCTTCCTGGAGACC 420  
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaLeuLysGluLeuLys 192  
Db 421 AGCGCCAAAGACTGGCATGATGAGTGGAGTGTAGCCTTTCTGCCATCGCAAGGAACTG 480  
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212  
Db 481 TACCGGCGCGGCATCAGCGGATGAGCCAGCTTCAGATCGGACATGTATGTAGATCC 540  
QY 213 GlnLysLysArgSerSerCysCysSerPheMet 223  
Db 541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATG 573  
RESULT 6  
BC007681  
LOCUS  
DEFINITION Homo sapiens, RAB26, member RAS oncogene family, clone MGC:3503  
IMAGE:3627067, mRNA, complete cds.  
ACCESSION BC007681  
VERSION BC007681.1 GI:14043378  
KEYWORDS MGC.  
SOURCE Homo sapiens.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1513)  
AUTHORS Strausberg R.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: http://www.nisc.nih.gov/  
Contact: nisc\_mgc@nri.nih.gov  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Image: 12 Row: m Column: 2  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5931611.  
FEATURES  
Location/Qualifiers  
1..1513  
/organism="Homo sapiens"  
/db\_xref="LocusID:25837"  
/db\_xref="taxon:9606"  
/clone="MGC:3503 IMAGE:3627067"  
/tissue\_type="uterus, endometrium adenocarcinoma"  
/clone\_lib="NIH\_MGC\_44"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
187..759  
/codon\_start=1  
/product="RAB26, member RAS oncogene family"  
/protein\_id="AAH07681.1"  
/db\_xref="GI:14043379"  
/translation="MLVGDGSGVGTCLLVRFKDGAFAGTGFISTVIGIDFRNKVLVDYGVKVKLQMDTAGOERFSVTHAYYRDHALLLYDVTNKASFDNICAWLTHEYAOHDVALMLGNKVDSAHERVVKREDGKLAKEYGLFPFMTSAKTGLNVDLAFATLAKELKQSRMKAPSEPRRLHDYVVKREGRGASCCRP"  
BASE COUNT 326 a 452 c 449 g 286 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,28e-70 Length: 1513  
Score: 780.50 Matches: 153  
Percent Similarity: 79.57% Conservatives: 30  
Best Local Similarity: 66.52% Mismatches: 32  
Query Match: 67.87% Indels: 16  
DB: 9 Gaps: 3  
US-09-817-199A-2 (1-223) x BC007681 (1-1513)  
QY 6 GlyAlaValAlaThrArgAspGlyGluAla-----ProGluArgSerPro----- 20  
Db 62 GGGGCGCGGCGGCGCTCCGGGACTCGCTTTCCGGCCCCCGCGCGCCACACGGG 121  
QY 21 ProCysSerProSer-----TyrAspLeuThrGly 30  
Db 122 CCTTGAGCGCGGC-CGGCCCTCGCTTGGCGGGGTGTCGACTTCTACGACGTGCGCTTC 180

```
QY 31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLys 50
Db 181 AAGGTCATGCTGGTGGGGAGCTGGGTGTGGGAAGACCTGCTGCTGGTGGCATCAAG 240
QY 51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
Db 241 GATGGTCTTCTCTGGGGGAGCTTCATCTCCACCGTAGGCATTGACTTCCCGGAACAA 300
QY 71 ValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
Db 301 GTTCTGACGTGATGTTGTGAGGTGAAGCTGCAGATGTGGACACAGCTGGTCAGGAG 360
QY 91 ArgPheArgSerValThrHisAlaTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyr 110
Db 361 CGGTTCGCGAGTGTACCCATGCTACTACCGGATGCTCATGCTGCTGCTGCTCTAC 420
QY 111 AspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
Db 421 GATGTACCAACAAAGGCTCTCTTTGACACATCCAGGCTGGCTGACCGAGATCCAGAG 480
QY 131 TyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
Db 481 TACCCCGACGACGAGTGGGCTCATGCTGCTGGGAACAGGTGGACTGTGCCATGAG 540
QY 151 ArgValIleArgSerGluAspGlyGluThrLeuAlaAlaArgGluTyrGlyValPropheLeu 170
Db 541 CGTGTGTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
Db 601 GAGACGAGCCCAAGAGGCGCTCAAGCTGGACTTGGCCCTTCACAGGCATAGCAAGAG 660
QY 191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
Db 661 TTGAGGAGCGCTCCATGAGGCTCCAGCGAGCGCGGCTTCCGGCTGCATGATTAGCT 720
QY 211 GluSerGlnLysLysArgSerSerCysCys 220
Db 721 AAGAGGAGGAGGTCGAGGCGGCTCCTGCTGC 750
```

```
RESULT 7
AB027137 1320 bp mRNA linear PRI 29-SEP-1999
LOCUS Homo sapiens v46133 mRNA for RAB-26, complete cds.
DEFINITION AB027137
ACCESSION AB027137
VERSION AB027137.1 GI:5931611
KEYWORDS RAB-26; v46133.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1320)
AUTHORS Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S.,
Muramatsu,M. and Saito,T.
TITLE Human RAS-related protein RAB-26
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 1320)
AUTHORS Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S.,
Muramatsu,M. and Saito,T.
TITLE Direct Submission
JOURNAL Radiological Sciences, Genome Research Group; Inage-ku Anagawa
4-9-1, Chiba, Chiba 263-8555, Japan (E-mail:t_saito@nirs.go.jp,
Tel:81-43-201-3135, Fax:81-43-251-9818)
FEATURES
source Location/Qualifiers
1..1320
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1320
/genes="v46133"
4..576
/genes="v46133"
```

```
/note="RAS-related protein"
/codon_start=1
/product="RAB-26"
/protein_id="BAA84707.1"
/db_xref="GI:5931612"
/translation="MLVGSQGVKTKLLVRFKDGAFAGTISTVGVIDPRNKVLDYDG
VKVKLQMDTQAGRSFVTHAYIRDAHALLLLYDVTNKASFDNIQAWLITEHEYAOH
DYVALLMLGNKVDKSAHERVYVYKREKLAKEYGLPFMETSAKTGLNLDLAATAKELK
ORSMKAPSEPRFLHDYVKREGRGASCCRP"
BASE COUNT 273 a 379 c 395 g 273 t
ORIGIN
```

## Alignment Scores:

Pred. No.:	2,78e-68	Length:	1320
Score:	757.00	Matches:	139
Percent Similarity:	88.36%	Conservative:	28
Best Local Similarity:	73.54%	Mismatches:	22
Query Match:	65.83%	Indels:	0
DB:	9	Gaps:	0

US-09-817-199A-2 (1-223) x AB027137 (1-1320)

QY	32	ValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAsp	51
Db	1	GTCTGCTGGTGGGGAGCTCGGTGTGGGAAGACCTGTCTGTGTGGTCAAGAT	60
QY	52	GlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysVal	71
Db	61	GGTGTCTTCTGGGGGAGCTTCATCTCCACCGTAGGCATTGATCCCGAACAAAGTT	120
QY	72	ValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArg	91
Db	121	CTGACGTGATGCTGTGAAGTGAAGTGCAGATGTGGACACAGCTGTCAGGAGCGG	180
QY	92	PheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAsp	111
Db	181	TTCGCGAGTGTACCCATGCTACTACCGGATGCTCATGCTGTGCTGCTCTACGAT	240
QY	112	IleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyr	131
Db	241	GTCAACAAAGGCTCTTTTGACACATCCAGGCTGGCTGACCGAGATCCAGGATAC	300
QY	132	AlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArg	151
Db	301	GCCAGCACGACGCTGGCGCTCATGCTGGGAACAAGGTGGACTGTGCCCATGAGCGT	360
QY	152	ValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValPropheLeuGlu	171
Db	361	GTGGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	420
QY	172	ThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeu	191
Db	421	ACAGGCCCAAGAGGCGCTCAAGTGGACTTGGCTTCACGCCATAGCAAGGAGTTG	480
QY	192	LysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGlu	211
Db	481	AAGCAGCGCTCCATGAAGGCTCCAGCGAGCGCGCTTCCGGCTGCATGATTAGCTTAAG	540
QY	212	SerGlnLysLysArgSerSerCysCys	220
Db	541	AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	567

## RESULT 8

RNUI8771

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Rattus norvegicus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

RNU18771 Rattus norvegicus Rab26 mRNA, complete cds. ROD 18-JUL-1995

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

1098 bp mRNA linear

```

REFERENCE
AUTHORS 1 (bases 1 to 1098)
TITLE Wagner,A.C., Strowski,M.Z., Goke,B. and Williams,J.A.
JOURNAL Molecular cloning of a new member of the Rab protein family, Rab
MEDLINE 26, from rat pancreas
PUBMED 95169156
BIOCHEM. Biophys. Res. Commun. 207 (3), 950-956 (1995)
7864900
REFERENCE 2 (bases 1 to 1098)
AUTHORS Williams,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1994) John A. Williams, Physiology, University of
Michigan, 7744 Med. Sci. II, Ann Arbor, MI 48109, USA
FEATURES
source
1..1098
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="female"
/tissue_type="pancreas"
/dev_stage="adult"
1..29
30..602
/codon_start=1
/product="Rab26"
/protein_id="AAA69955.1"
/db_xref="GI:619734"
/translation="MLVGDGSGVGTCLLVRFKDGAGLAGTFTSTVGIDFRNKVLVDVG
MKVKLQIWDTAGQERFRSVTHAYYRDAHALLLYDITNKDSFDNIQAWLTIHEYAOQ
DVLMLLGNKVDSTQERVKREDGEKLAKEYGLPFMETSAKSLGNLVDLAFIAKELK
QRSTKAPSEPRFLHDYVKREGVSCRL"
603..1098
3'UTR
polyA_signal 1063..1068
BASE COUNT 267 a 277 c 306 g 248 t
ORIGIN
Alignment Scores:
Pred. No.: 4.44e-68 Length: 1098
Score: 754.00 Matches: 140
Percent Similarity: 87.89% Conservative: 27
Best Local Similarity: 73.68% Mismatches: 23
Query Match: 65.57% Indels: 0
DB: 10 Gaps: 0

US-09-817-199A-2 (1-223) x RNUI8771 (1-1098)
Qy 31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLys 50
Db 24 AAGGTCATGCTGGTGGGGAGTCCGGTGTGGGAAGACGTGCTGCTCGCTTCAAG 83
Qy 51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
Db 84 GATGGGGCTTCCTGGCTGGTACCTTCATCTCCACTGTGGGATCGACTTCGGGAATAA 143
Qy 71 ValValThrValAspGlyValargValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
Db 144 GTTCTGGATGGTGGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 203
Qy 91 ArgPheArgSerValThrHisAlaThrTyrrArgaspAlaGlnAlaLeuLeuLeuTyr 110
Db 204 CGTTCGCAAGTGTACCCATGCCCTACTACCGGGATGTCCAGCAGTGTGCTGCTAC 263
Qy 111 AspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
Db 264 GACATCACCAACAAGATTCCTTCGACAAACATCCAGGCTGGTGGACAGAATCCAGAA 323
Qy 131 TyrAlaGlnArgaspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
Db 324 TATGCCACAGCAGCGTGTGCTCATGCTGTGGGAACAAGTTGACTTACTCAAGAA 383
Qy 151 ArgValIleArgSerGluAspGlyCluThrLeuAlaArgGluTyrGlyValPropheLeu 170
Db 384 CGAGTGGTAAAGAGAGATGGGGAGAAATTAGCCAAAGGAGATATGGGCTGCCATTCATG 443

```

```

Qy 171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
Db 444 GAGACACGCGCAAGAGCGGCTCAATGTGGACTTGGCTTTTACAGCCATACCAAGGAG 503
Qy 191 LeuLysTyrArgAlaGlyHisGlnAlaaspGluProSerPheGlnIleArgaspTyrVal 210
Db 504 CTGAACAACAAGATCCACCAAGGCTCCAGTAGCGCGCTTCAGGCTGCATGACTATGTG 563
Qy 211 GluSerGlnLysLysArgSerSerCysCys 220
Db 564 AAGAGGAGGCGGCGGAGGCTCTCTCTGTGT 593
RESULT 9
AF498952
LOCUS Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete
DEFINITION AF498952.1 GI:20379079
VERSION AF498952
KEYWORDS CDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
TITLE Homo sapiens RAB family small GTP binding protein RAB26
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 573)
AUTHORS Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA
FEATURES
source
1..573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
1..573
/gene="RAB26"
1..573
/gene="RAB26"
/feature="RAB family member"
/codon_start=1
/product="small GTP binding protein RAB26"
/protein_id="AAM21100.1"
/db_xref="GI:20379080"
/translation="MLVGDGSGVGTCLLVRFKDGAGLAGTFTSTVGIDFRNKVLVDVG
VKVKLQIWDTAGQERFRSVTHAYYRDAHALLLYDITNKDSFDNIQAWLTIHEYAOQ
DVLMLLGNKVDSTQERVKREDGEKLAKEYGLPFMETSAKSLGNLVDLAFIAKELK
QRSTKAPSEPRFLHDYVKREGVSCRL"
BASE COUNT 117 a 151 c 191 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 2.41e-68 Length: 573
Score: 753.00 Matches: 138
Percent Similarity: 88.30% Conservative: 28
Best Local Similarity: 73.40% Mismatches: 22
Query Match: 65.48% Indels: 0
DB: 9 Gaps: 0

US-09-817-199A-2 (1-223) x AF498952 (1-573)
Qy 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysaspGly 52
Db 1 ATGCTGGTGGGGAGTCCGGTGTGGGAAGACCTGCTGCTGCTGCGATTCAAGGATGT 60
Qy 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
Db 61 GCTTTCCTGGCGGGAGACCTTCATCTCCACCGGTAGCATTTGACTTCCCGAACAAGTTCGT 120

```

```

QY 73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
Db 121 GACGTGATGTTGAGAGTGAAGCTGCATGTTGGACACAGCTGTCAGGAGCGGTT 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
Db 181 CGCAGTGTACCATGCTTACTACCGGATGCTCATGCTCTGCTGCTCTACGATGC 240
QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
Db 241 ACCAACAGGCCCTTTTGACACATCCAGCGCTGCTGACCCAGATCCACGAGTACGCC 300
QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
Db 301 CAGCACAGCTGCGCTCATGCTGCTGGGGAACAAGTGGACTCTGCCATGAGCGTGT 360
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 361 GTGAAGAGGAGGAGCGGGAGAGCTGGCCACAGGATGTAGACTGCCCTTCATGGAGACC 420
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 421 AGCGCAAGACGGGCTCAACGTGGACTTGGCCCTTCACAGCCATAGCAAGGAGTTGAAG 480
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 481 CAGCGTCCATGAAGCTCCACGAGCGCGCTTCGCGTGCATGATTACGTTAAGAGG 540
QY 213 GlnLysLysArgSerSerCysCys 220
Db 541 GAGGTCGAGGGGCTCTGCTGC 564

RESULT 10
AR062279
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

AR062279
Sequence 2 from patent US 5843717.
AR062279
AR062279.1 GI:5989970
Unknown.
Unclassified.
Hillman,J.L. and Guegler,K.J.
Rab protein
Patent: US 5843717-A 2 01-DEC-1998;
Location/Qualifiers
1..1340
/organism="unknown"
276 a 387 c 403 g 274 t

Alignment Scores:
Pred. No.: 1,32e-65 Length: 1340
Score: 731.00 Matches: 141
Percent Similarity: 86.73% Conservative: 29
Best Local Similarity: 71.94% Mismatches: 25
Query Match: 63.57% Indels: 2
DB: 6 Gaps: 0

US-09-817-199a-2 (1-223) x AR062279 (1-1340)

QY 26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
Db 22 TAGCAGCTGCCCTTCAAGTTCATGCTGGTGGGAGCTCGGGTGTGGGAAGACCTGCTG 81
QY 46 Leu-IleGlnPheLysAspGlyValAlaPheLeuSerGluThrPheIleAlaThrValClytl 65
Db 82 CTGGGTGGCATTCAGGATGGTGTCTTCTGGCGGGAGCCTTCATCTCCACCGTAGC-AT 140
QY 65 eAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAs 85
Db 141 TGACTTCCGGAACAAAGTTCTGGACGTGGATGTTGTGAAGGTGAAGCTGCACATGTGGGA 200

```

```

QY 85 pThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAl 105
Db 201 CACAGCTGGTCAGGAGCGGTTCCGCACTGTTACCCATGCTACTACCGGATGCTCATGC 260
QY 105 aLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLe 125
Db 261 TCTGCTGCTGCTACGATGTACCAACAAGGCTCTCTTTGACAAACATCCAGGCTGGCT 320
QY 125 uThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAl 145
Db 321 GACCGAGATCCAGGATACGCCACGACGCTGGCGCTCATGCTGCTGGGGAACAAGGT 380
QY 145 aAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTy 165
Db 381 GGACTCTGCCCTGACGCTGTGGTGAAGAGGAGGAGCGGAGAGCTGCCAAGAGTA 440
QY 165 rGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLe 185
Db 441 TGGACTGCGCTTTCATGGAGACCGAGCGCAAGCGGCTCAACGTGGACTTGGCTTCAC 500
QY 185 uAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGl 205
Db 501 AGCCATAGCAAGAGGTTGAACGAGCGCTCCATGAAGGCTCCAGCGGCGGCTCCG 560
QY 205 nIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 561 GCTGCATGATTACGTTAAGAGGAGGCTGAGGCGGCTCTCTGCTGC 606

RESULT 11
AK054846
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

AK054846
3105 bp mRNA linear PRI 01-AUG-2002
Homo sapiens cDNA FLJ30284 fis, clone BRACE2002812, moderately
similar to Mus musculus GTPase Rab37 (Rab37) mRNA.
AK054846
AK054846.1 GI:16549463
oligo capping; fis (full insert sequence).
Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
clone:BRACE2002812.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,T., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3105)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
Location/Qualifiers
1..3105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2002812"
/tissue_type="cerebellum"
/clone_lib="BRACE2"

```



/note="cloning vector: pME18SFL3"

BASE COUNT	680 a	905 c	898 g	622 t
ORIGIN				
Alignment Scores:				
Pred. No.:	5e-63	Length:	3105	
Score:	710.50	Matches:	182	
Percent Similarity:	39.47%	Conservative:	13	
Best Local Similarity:	36.84%	Mismatches:	26	
Query Match:	61.78%	Indels:	274	
DB:	9	Gaps:	5	
US-09-817-199A-2 (1-223) x AK054846 (1-3105)				
QY	3	GlyThrProGlyValAlaIleAlaThrArgAspGlyGluAlaProGluArgSerProPro	--- 21	
DB	375	GGGAGGCGGAGCGCGAGCTCCAAAGCTGGCAGTGGACCTGCAGAGACCCGATTCCTAC	434	
QY	22	-----CysSerProSerTyrAsp-----LeuThrGlyLysValMetLeuLeuGly	36	
DB	435	CAGGGAGGAGCTGGCCCTGACTCAACGACACCGTCTGCATTAAGACCATCTCTGGTGGCT	494	
QY	37	AspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer	56	
DB	495	GACAGTGGTGGGAAAGACGCTCTGCTGGTTCAGTTCGATCAGGGCAAGTTCATCCCC	554	
QY	57	GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValIleValThrValAspGly	76	
DB	555	GGCTCTCTCGGCCACTGTGGCATCGGATTCAAGCAAGGTGGTACTGTGGATGGC	614	
QY	77	ValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThr	96	
DB	615	GTGAGAGTCAAGCTGCAGATCTGGACACCGCTGGCAGCAAGCGTTCCGAAGCGTCACC	674	
QY	97	HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSer	116	
DB	675	CATGCTTAFTACAGAGATGCTCAGGCGCTTGCCTGTATGACATCACCAACAATCT	734	
QY	117	SerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspVal	136	
DB	735	TCTTTGACACATCAGGGCGCTGGCTCACTGAGATTCATGATATGCCAGAGGACGTG	794	
QY	137	ValIleMetLeuLeuGlyAsnLys-----	144	
DB	795	GTGATCATGCTGTAGGCAACAAGGTGAGTGGTCCGGGCGAGGTGACGCCCGCTGC	854	
QY	144	-----	144	
DB	855	ACTTCTCAGCCCTAGCCGGCCCCCATAACCCCAAGACAGTTATCTAGGCATCCTTCC	914	
QY	144	-----	144	
DB	915	TGAAAGGACTCTGCAGCTCCAGCTCAGGCTCAGGGGTACACATATCTGGAGCTTCTGCCCAT	974	
QY	144	-----	144	
DB	975	CCCATCTGCCCTTCCAGGAAAGTCCAAGTTGTTGCTTGAGAAATCAAGGGGTGCCAG	1034	
QY	144	-----	144	
DB	1035	TTCTCAGCCCCCATTAGAGCAGAGTGAACAGGGTCCCGAGGTCCAGGGCTAAGAGTGCATA	1094	
QY	144	-----	144	
DB	1095	GGTTAGCCCCCACTGCTGCTCTATTCCAAAGACCCCTTTACCAAGGTGAGATCCCGAGC	1154	
QY	144	-----	144	
DB	1155	TGGAGGTACTGGGAGAAACCCCTGGCCCCGAGCCCAATCACACCTGCTGCAGTCCCT	1214	
QY	144	-----	144	
DB	1215	TGGGCCACCCAGAGGGCAGGCAACGGCTGCTTCTGGGGCAAAATATGGGCCGCTGGG	1274	

QY	145	-----	AlaaspMetSerSerG1	150		
Db	1275	GGCGAGCGCTCTCTCCACAGATGACCCATTGGGCTTGACAGCGGATATGACGACGGA	1334			
QY	150	uATGValIleArgSerGluAspGlyGluThrLeuAlaArg-----	163			
Db	1335	AAGAGTATCCGTTCCGAAGCAGGAGACCTTGGCCAGGGTAAAGTGAATGTGTGTGGGA	1394			
QY	164	-----	-G1 164			
Db	1395	CAGGGTGAAGGTGGGGCAACCCGACGCTGGCCCTGAGGACACTCTCTCCCGGGCAGGA	1454			
QY	164	uTYrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValIleLeuAlaPhe	184			
Db	1455	GTACGGTGTTCCTCTCTGGAGACCAAGCGCAAGACTGGCATGAATGTGGAGTTAGCCTT	1514			
QY	184	eLeuAlaIleAlaLys-----	189			
Db	1515	TCTGGGCATCGCCAA--GTGAGACTGGCGAGGAAGTGTGGGGCGAGGCGGCA	1573			
QY	189	-----	189			
Db	1574	CACTCCAGGAATCCAGTAGGGCCCGCCCTGGCCCGCCCTGGACACACCTGCATTCT	1633			
QY	189	-----	189			
Db	1634	GCAGGCTGAGGTCCATTGCTCTGGGAGCAGCTGGGCCACTGGGAGAGGGGGGGCGGC	1693			
QY	190	-----	-G1 190			
Db	1694	TCAGCTCTCTACCCCGCCAGCCAGCCAGCCAGCCAGCCATTTGCTCTTTCAGGGA	1753			
QY	190	uLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVa	210			
Db	1754	ACTGAATACCGGGCGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGT	1813			
QY	210	lGluSerGlnLysLysArgSerSerCysSerPheMet	223			
Db	1814	AGAGTCCAGAGAAGCGCTCCAGCTGCTGCTCTTCATG	1853			
RESULT	12					
LOCUS	AY061826	2536 bp	mRNA	linear		
DEFINITION	Drosophila melanogaster GH21984 full length cDNA.					
ACCESSION	AY061826					
VERSION	AY061826.1	GI:16902019				
KEYWORDS	FLI_CDNA.					
SOURCE	Drosophila melanogaster.					
ORGANISM	Drosophila melanogaster					
REFERENCE						
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.					
TITLE	1 (bases 1 to 2536)					
JOURNAL	Direct Submission					
COMMENT	Submitted (02-NOV-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory Berkeley, CA 94720 This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are					



/protein\_id="AAH19990.1"  
/db\_xref="GI:18043409"  
/translation="MAKTVDFYLLKLLIGDSGVGKTCVLFRESEDAFNSTFTSTIGID  
FKIRIEIDGKRILQIWDTAGOERFRTITAYXRGAMGIMLVYDITNEKSFSDNIRNW  
IRNIEHASADVEKMLGNKCDVNRKQVSKERGEKALDYGKFMETSAKANINVEN  
AFFTIARDIKAMDKKLENSPGSGVKITVEOKRTSFRCSSL"  
BASE COUNT 367 a 338 c 341 g 291 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.19e-44 Length: 1337  
Score: 521.50 Matches: 101  
Percent Similarity: 69.76% Conservative: 42  
Best Local Similarity: 49.27% Mismatches: 55  
Query Match: 45.35% Indels: 7  
DB: 10 Gaps: 3  
US-09-817-199a-2 (1-223) x BC019990 (1-1337)  
QY 25 SerTyAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCys 44  
Db 10 ACCTACGATTACCTGTTCAGCTGCTGCTGATCGGGGACTCGGGGTAGGAACCTGT 69  
QY 45 PheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGly 64  
Db 70 GTCTCTCTCCGCTTCCTCCGAGGACGCTTC---AACTCCACATTCATCTCTACCATAGGA 126  
QY 65 IleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTyr 84  
Db 127 ATGACTTAAATATTAGGACCATAGAGCTCGATGGCAAGAGATTAACTGCAGATG 186  
QY 85 AspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGln 104  
Db 187 GACACGGCCGCGCAGGAGCGGTTCGAACAATCATCACGACGCTACTACAGGGGTGCCATG 246  
QY 105 AlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTyr 124  
Db 247 GGTATCATGCTGTCTACGACATATTACCAATGAGAGCTCTTTGACAACATCCGGAATTGG 306  
QY 125 LeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLys 144  
Db 307 ATTCGGAACATTGAAGAGATGCTCTCGAGAGCTGGAGAGATGATCTGGGGAAATAG 366  
QY 145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164  
Db 367 TGTGATGTGAATGACAAAGACAGGTGTCCAAAGGACGGGAGAAAGCTGGCACTCGAC 426  
QY 165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184  
Db 427 TATGGATCAAGTTCATGGAGACCACTGCAAGGCCCAACATTAATGTGAGAATGCATTT 486  
QY 185 LeuAlaIleAlaLysGluLeuLysTyrArg-----AlaGlyHisGlnAla 199  
Db 487 TTCACCTCTTCCAGGAGATATCAAGCAAAATGAGCAAAATGGAAGGGAACACCCCG 546  
QY 200 AspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCys 219  
Db 547 CAGGGGAGCAGCCATGAGTCAAGATCAGATGAGCAGCAGAGAGAGACCAGCTTCTTC 606  
QY 220 ---CysSerPheMet 223  
Db 607 CGGTGCGAGTCTCTCTG 621

RESULT 14  
AB024994  
LOCUS  
DEFINITION  
AB024994  
Cicer arietinum mRNA for rab-type small GTP-binding protein,  
complete cds.  
AB024994  
ACCESSION  
AB024994.1 GI:4586579  
VERSION  
KEYWORDS  
rab-type small GTP-binding protein.  
Cicer arietinum (strain:ILC3279) leaf cDNA to mRNA, clone:INR134.  
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;  
Cicer.  
1 (sites)  
Ichinose,Y., Tiemann,K., Schwenger-Erger,C., Toyoda,K., Hein,F.,  
Hanselle,T. and Barz,W.  
Genes Expressed in Aschochyta rabiei-Inoculated-Chickpea Plants and  
Elicited Cell Cultures as Detected by Differential  
cDNA-Hybridization  
Unpublished  
REFERENCE  
2 (bases 1 to 1074)  
AUTHORS  
Ichinose,Y.  
TITLE  
Submitted (16-MAR-1999) Yuki Ichinose, Okayama University, Faculty  
of Agriculture, Tsushima-naka, 1-1-1, Okayama, Okayama 700-8530,  
Japan (E-mail:yuki@cc.okayama-u.ac.jp, tel: +81-86-251-8308,  
Fax: +81-86-251-8308)  
FEATURES  
Location/Qualifiers  
1..1074  
/organism="Cicer arietinum"  
/strain="ILC3279"  
/db\_xref="taxon:3827"  
/clone="INR134"  
/tissue\_type="leaf"  
107..715  
/codon\_start=1  
/product="rab-type small GTP-binding protein"  
/protein\_id="BAH76422.1"  
/db\_xref="GI:4586580"  
/translation="MNPEYDYLFLKLLIGDSGVGKSCLLLRADDSYLDYSISTIGVD  
FKIRVEQDGKTKIQIWDTAGOERFRTITSSYRGAGHIIIVYDVTQDESNNVQKW  
LNEIDRYASENVNKLGVNKCDAANKVSVSSAKAFADIEIGIPEMETSAKNATVEQ  
APMAAAEIKNEMASQPNANRPPTVQIRGQPMNQKSCCSI"  
1074  
polyA\_site  
/note="8 A nucleotides"  
BASE COUNT 310 a 223 c 207 g 334 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6.4e-44 Length: 1074  
Score: 518.50 Matches: 97  
Percent Similarity: 69.90% Conservative: 47  
Best Local Similarity: 47.09% Mismatches: 55  
Query Match: 45.09% Indels: 7  
DB: 8 Gaps: 3  
US-09-817-199a-2 (1-223) x AB024994 (1-1074)  
QY 20 ProProCysSerProSerTyAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39  
Db 101 CCCGTATGATCCCAATATGACTATTCTTCAAGCTTTGTGATGGAGATCTCGT 160  
QY 40 ValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59  
Db 161 GTGGCAAGTCATGCTCTCTCTGAGTGTGCTGATGATTCATACCTGAC---AGCTAT 217  
QY 60 IleAlaThrValGlyLysPheArgAsnLysValValThrValAspGlyValArgVal 79  
Db 218 ATCAGTACAAATGGAGTGGACTTTAAATTCGCACGTGTTGAGCAAGACGGAAGACCAT 277  
QY 80 LysLeuGlnIleTyrAspThrAlaGlyGlnArgPheArgSerValThrHisAlaTyr 99  
Db 278 AACTTCAATTTGGGACACTGCTGTCAGAACGTTTCGGGACTATCTACTACGACTAC 337  
QY 100 TyrArgAspAlaGlnAlaLeuLeuLeuLeuLeuTyAspIleThrAsnLysSerPheAsp 119  
Db 338 TATCGTGGGCTCATGGCATAATTGTTTATGATGTCACCTACCAAGAGAGCTTTAAC 397  
QY 120 AsnIleArgAlaThrPheLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMet 139  
Db 398 AATGTTAAGCAGTGGCTGAATGAATTCGCGTATGCAAGTGAATAATGTAACAAGCTT 457



Result No.	Query	Score	Match	Length	DB	ID	Description
1	1150	100.0	223	1	RB37_HUMAN	Q96ax2	homo sapien
2	1081	94.0	223	1	RB37_MOUSE	Q91km7	mus musculus
3	753	65.5	190	1	RB26_HUMAN	Q9ulw5	homo sapien
4	745	64.8	190	1	RB26_RAT	P51156	rattus norv
5	517.5	45.0	203	1	SAS2_DICDI	P20791	dictyosteli
6	516	44.9	206	1	RAB8_MOUSE	P55258	mus musculus
7	513.5	44.7	207	1	RAB8_HUMAN	P24407	homo sapien
8	512.5	44.6	215	1	RAB1_BETUO	Q39433	beta vulgar
9	512	44.5	210	1	RAB8_DISOM	P21218	discopge o
10	507	44.1	258	1	ARA5_ARATH	P28188	arabidopsis
11	502.5	43.7	207	1	RB8B_RAT	P70550	rattus norv
12	501	43.6	207	1	RB8B_HUMAN	Q29230	homo sapien
13	500	43.5	202	1	RIC1_ORYSA	P40392	oryza sativ
14	498	43.3	208	1	SAS1_DICDI	P20790	dictyosteli
15	495	43.0	203	1	YPR2_WAIZE	Q05737	zea mays (m
16	495	43.0	216	1	ARA3_ARATH	P28186	arabidopsis
17	494	43.0	200	1	RB10_CANFA	P24409	canis famli
18	494	43.0	200	1	RB10_HUMAN	O88386	homo sapien
19	490.5	42.7	200	1	RAO1_DISOM	P22127	discopge o
20	487	42.3	200	1	YPR2_SCHPO	P17609	schizosacch
21	486	42.3	203	1	YPR1_CHLRE	Q39571	chlamydomon
22	479.5	41.7	203	1	YPR1_WAIZE	P16976	zea mays (m
23	478	41.6	203	1	YPR1_VOLCA	P33584	volvox cart
24	477.5	41.5	205	1	RAB1_LYNST	Q05974	lymnaea sta
25	476	41.4	217	1	YPR2_VOLCA	P36861	volvox cart
26	473.5	41.2	203	1	RB13_HUMAN	P51153	homo sapien
27	472	41.0	201	1	RB1B_RAT	P10536	rattus norv
28	465.5	40.5	202	1	RAB1_DISOM	P22125	discopge o
29	464.5	40.4	203	1	RYL1_YARLI	P41924	yarrowia li
30	461	40.1	200	1	RB10_RAT	P35281	rattus norv
31	459.5	40.0	203	1	YPR1_NEUCR	P33723	neurospora
32	458.5	39.9	205	1	RB1A_HUMAN	P11476	homo sapien
33	456.5	39.7	205	1	RB1A_RAT	P05711	rattus norv

```
QY 121 IRAWLTHEYAQDVVIMLGNKADMSSEVRSRSDGETLAREYGVPELETSAGTGMV 180
|||||
Db 121 IRAWLTHEYAQDVVIMLGNKADMSSEVRSRSDGETLAREYGVPELETSAGTGMV 180
|||||
QY 181 ELAFLAIAKELKYRAGHQADSFQIRDVESQKRSSCCSF 223
|||||
Db 181 ELAFLAIAKELKYRAGHQADSFQIRDVESQKRSSCCSF 223
|||||

RESULT 2
RB37_MOUSE
ID RB37_MOUSE STANDARD; PRT; 223 AA.
AC Q9JRM7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-37.
GN RAB37.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20189834; PubMed=10722846;
RA Masuda E.S., Luo Y., Young C., Shen M., Rossi A.B., Huang B.C., Yu S.,
RA Bennett M.K., Pavan D.G., Scheller R.H.;
RT "Rab37 is a novel mast cell specific GTPase localized to secretory
RT granules";
RL FEBS Lett. 470:61-64(2000).
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: Specifically expressed in the bone marrow mast
CC cells.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF233582; AAF67162.1; -.
DR HSSP; P01112; 1PLJ.
DR MGD; MGI:1929945; Rab37.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Protein transport.
FT NP_BIND 36 43 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 143 146 GTP (BY SIMILARITY).
FT DOMAIN 59 67 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 219 219 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 220 220 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 223 AA; 24716 MW; E72013621C12C605 CRC64;

Query Match 94.0%; Score 1081; DB 1; Length 223;
Best Local Similarity 93.7%; Pred. No. 2.9e-88;
Matches 209; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MTGTPGAVATRGCEAPERSPPSPVDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTPI 60
|||||
Db 1 MTGTPGAATAGDGEAPERSPPSPVDLTGKVMLLGDSGVGKTCFLIQKDGAFLSGTPI 60
|||||

QY 61 ATVGIDFRNKVTVDCGVRYKLOIWDTAGOERFSVTHAYYRDAQALLLYDITNKSSFDN 120
|||||
Db 61 ATVGIDFRNKVTVDCGVRYKLOIWDTAGOERFSVTHAYYRDAQALLLYDITNKSSFDN 120
|||||
```

```
QY 121 IRAWLTHEYAQDVVIMLGNKADMSSEVRSRSDGETLAREYGVPELETSAGTGMV 180
|||||
Db 121 IRAWLTHEYAQDVVIMLGNKADMSSEVRSRSDGETLAREYGVPELETSAGTGMV 180
|||||
QY 181 ELAFLAIAKELKYRAGHQADSFQIRDVESQKRSSCCSF 223
|||||
Db 181 ELAFLAIAKELKYRAGHQADSFQIRDVESQKRSSCCSF 223
|||||

RESULT 3
RB26_HUMAN
ID RB26_HUMAN STANDARD; PRT; 190 AA.
AC Q9ULW5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-26.
GN RAB26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal brain;
RX MEDLINE=20496221; PubMed=11043516;
RA Seki N., Yoshikawa T., Hattori A., Miyajima N., Muramatsu M.,
RA Saito T.;
RT "cDNA cloning of a human RAB26-related gene encoding a Ras-like
RT GTP-binding protein on chromosome 16p13.3 region.";
RL J. Hum. Genet. 45:309-314(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB027137; BAA84707.1; -.
DR EMBL; BC007681; AAH07681.1; -.
DR HSSP; P36017; 1ER0.
DR Genew; HGNC:14259; RAB26.
DR MIM; 605455; -.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation.
FT NP_BIND 4 11 GTP (BY SIMILARITY).
FT NP_BIND 53 57 GTP (BY SIMILARITY).
FT NP_BIND 111 114 GTP (BY SIMILARITY).
FT DOMAIN 27 35 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 187 187 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 188 188 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 190 AA; 21354 MW; 517387B7C0B3291C CRC64;

Query Match 65.5%; Score 753; DB 1; Length 190;
Best Local Similarity 73.4%; Pred. No. 2e-59;
Matches 138; Conservative 28; Mismatches 22; Indels 0; Gaps 0;
```

```

QY 33 MLGDTGVGKTCFLIQKDGAFSLGTFIATVGIIDFRNKVYVVDGVVRVKLIQIWDTAGQERF 92
Db 1 MLVGDGSGVGTCLLVRFKDGAFSLGTFIATVGIIDFRNKVYVVDGVVRVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEHYAQRDQVVMILGNKADMSSEV 152
Db 61 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEHYAQRDQVVMILGNKADMSSEV 120
QY 153 IRSEDEGTAREYGVPFLETSKTMNVLAFATIAKELKYRAGHOADPSFOIRDYVES 212
Db 121 VREDEGEKLAKEYGLPFMETSAKSLGNVLDLAFATIAKELKQSRKAPSEPRFLHDYVKR 180
QY 213 QKRRSSCC 220
Db 181 EGRGASCC 188

RESULT 4
RB26_RAT
ID RB26_RAT STANDARD; PRT; 190 AA.
AC P51156;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Ras-related protein Rab-26.
GN RAB26.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95169156; PubMed=7864900;
RA Wagner A.C.C., Sirovski M.Z., Goetz B., Williams J.A.;
RT "Molecular cloning of a new member of the Rab protein family, Rab 26,
RT from rat pancreas.";
RL Biochem. Biophys. Res. Commun. 207:950-956(1995).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PANCREAS, KIDNEY, BRAIN,
CC SUBMANDIBULAR GLAND, AND LUNG.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18771; AAA69955.1; -
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR Pfam; PF00071; ras; 1.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation.
FT NP_BIND 4 11 GTP (BY SIMILARITY).
FT NP_BIND 53 57 GTP (BY SIMILARITY).
FT NP_BIND 111 114 GTP (BY SIMILARITY).
FT DOMAIN 27 35 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 187 187 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 188 188 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 190 AA; 21457 MW; 96B2D1661F363031 CRC64;

Query Match 64.8%; Score 745; DB 1; Length 190;
Best Local Similarity 73.4%; Pred. No. 1e-58;
Matches 138; Conservative 27; Mismatches 23; Indels 0; Gaps 0;

QY 33 MLGDTGVGKTCFLIQKDGAFSLGTFIATVGIIDFRNKVYVVDGVVRVKLIQIWDTAGQERF 92
Db 1 MLVGDGSGVGTCLLVRFKDGAFSLGTFIATVGIIDFRNKVYVVDGVVRVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEHYAQRDQVVMILGNKADMSSEV 152
Db 61 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEHYAQRDQVVMILGNKADMSSEV 120
QY 153 IRSEDEGTAREYGVPFLETSKTMNVLAFATIAKELKYRAGHOADPSFOIRDYVES 212
Db 121 VREDEGEKLAKEYGLPFMETSAKSLGNVLDLAFATIAKELKQSRKAPSEPRFLHDYVKR 180
QY 213 QKRRSSCC 220
Db 181 EGRGASCC 188

```

```

Db 1 MLVGDGSGVGTCLLVRFKDGAFSLGTFIATVGIIDFRNKVYVVDGVVRVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEHYAQRDQVVMILGNKADMSSEV 152
Db 61 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLTIHEHYAQRDQVVMILGNKADMSSEV 120
QY 153 IRSEDEGTAREYGVPFLETSKTMNVLAFATIAKELKYRAGHOADPSFOIRDYVES 212
Db 121 VREDEGEKLAKEYGLPFMETSAKSLGNVLDLAFATIAKELKQSRKAPSEPRFLHDYVKR 180
QY 213 QKRRSSCC 220
Db 181 EGRGASCC 188

RESULT 5
SAS2_DICDI
ID SAS2_DICDI STANDARD; PRT; 203 AA.
AC P20791;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GTP-binding protein SAS2.
GN SASB OR SAS2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220623; PubMed=2109188;
RA Saxe S.A., Kimmel A.R.;
RT "SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum
RT with sequence similarities to essential genes in Saccharomyces
RT cerevisiae.";
RL Mol. Cell. Biol. 10:2367-2378(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89209367; PubMed=3149563;
RA Saxe S.A., Kimmel A.R.;
RT "Genes encoding novel GTP-binding proteins in Dictyostelium.";
RL Dev. Genet. 9:259-265(1988).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34456; AAA33248.1; -
DR PIR; B34716; B34716.
DR PIR; B61571; B61571.
DR HSSP; P05713; 3RAB.
DR DictyDb; DD05021; sasB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Palmitate.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT LIPID 202 202 PALMITATE (BY SIMILARITY).
FT LIPID 203 203 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 203 AA; 22642 MW; 33BFB2F451ECAC96 CRC64;

```

Query Match 45.08; Score 517.5; DB 1; Length 203;  
Best Local Similarity 47.18; Pred. No. 1.3e-38;  
Matches 97; Conservative 50; Mismatches 54; Indels 5; Gaps 3;

QY 15 APERSPPSPSYDITGKVMLLGDTGKTCFLIQKDGAFSLGTFIATVGIDFRNKVVTV 74  
DB 3 SPATNKA--AYDFLVKLLIGSGVSKSLLRFSQSP-TPSFIATIGIDFKIRITEL 59  
QY 75 DGVVKLIQIWDTAGQERFRSVTHAYYRDAQALLLYDITNKSFDNIRAWLTIHEYAQR 134  
DB 60 EGKRIKLIQIWDTAGQERFRITAYYRGAMGILLVYDTEKSFGRNWRNIEQHASD 119  
QY 135 DVVIMLGNKADMSRVSIRSEGETIAREYGVFFLETSAGTGMNVELAFALAKELKYR 194  
DB 120 SVNKLGNKCDMTKRVKVDSSAKSLADEYGIKFLTSAKSNVYVEAFIGLAKDIKKR 179  
QY 195 AGHQADEPFIQIDYVESQKKRSSCC 220  
DB 180 MIDTPNDPDHTI--CITPNKNKNTCC 203

## RESULT 6

RAB8\_MOUSE  
ID RAB8\_MOUSE STANDARD; PRT; 206 AA.  
AC P55258;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Ras-related protein Rab-8 (Oncogene c-mel).  
GN RAB8 OR MEL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91360267; PubMed=1886711;  
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,  
RA Johnson K.J.;  
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related  
RT genes";  
RL Oncogene 6:1347-1351(1991).  
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND  
CC NEUROTRANSMITTER RELEASE.  
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DB EMBL; S53270; AAB19682.1; -  
DB HSSP; P05713; 3RAB.  
DB MGD; MGI:96960; Mel.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001230; Prenyl\_site.  
DR InterPro; IPR001806; Ras\_trnsfrmng.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTENFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRfams; TIGR00231; small\_GTP; 1.  
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;  
KW Proto-oncogene.  
FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).  
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D662942 CRC64;

Query Match 44.9%; Score 516; DB 1; Length 206;  
Best Local Similarity 49.5%; Pred. No. 1.7e-38;  
Matches 101; Conservative 41; Mismatches 56; Indels 6; Gaps 3;

QY 25 SYDLTKVMLLGDTGKTCFLIQKDGAFSLGTFIATVGIDFRNKVVTVGVRVKLIQIW 84  
DB 4 TYDLYFKLLIGSGVSKSLLRFSQSP-NSTFISTIGIDFKIRITELDGKRIKLIQIW 62  
QY 85 DTAGQERFRSVTHAYYRDAQALLLYDITNKSFDNIRAWLTIHEYAQRDVVIMLGNK 144  
DB 63 DTAGQERFRITAYYRGAMGIMLVYDITNEKSFNIRNWRNIEEHASADVEKMILGNK 122  
QY 145 ADMSRVSIRSEGETIAREYGVFFLETSAGTGMNVELAFALAKELKYRAGHQ---AD 200  
DB 123 CDVNDKRVKSGEKGKALDYGKFMETSAKANINVENAFITLARDIKAKMDKNWKATAA 182  
QY 201 EPSFQIRDYVESQKKRSSC-CSEFM 223  
DB 183 GSSHGKTIIVEQOKRTSFRCSSL 206

## RESULT 7

RAB8\_HUMAN  
ID RAB8\_HUMAN STANDARD; PRT; 207 AA.  
AC P24407;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Ras-related protein Rab-8 (Oncogene c-mel).  
GN RAB8 OR MEL.  
OS Homo sapiens (Human), and  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606, 9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RX MEDLINE=94124602; PubMed=8294494;  
RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,  
RA Tavittian A., Louvard D.;  
RT "A small rab GTPase is distributed in cytoplasmic vesicles in non  
RT polarized cells but colocalizes with the tight junction marker ZO-1  
RT in polarized epithelial cells";  
RL J. Cell Biol. 124:101-115(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RX MEDLINE=91360267; PubMed=1886711;  
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,  
RA Johnson K.J.;  
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related  
RT genes";  
RL Oncogene 6:1347-1351(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human; TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=C.familiaris; STRAIN=Cocker spaniel;  
RX MEDLINE=91061765; PubMed=2123294;  
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;  
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell  
RT line";  
RL Mol. Cell. Biol. 10:6578-6585(1990).  
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND  
CC NEUROTRANSMITTER RELEASE.  
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X56741; CAA40065.1; -  
DR EMBL; S53268; AAB19681.1; -  
DR EMBL; BC002977; AAH02977.1; -  
DR EMBL; X56385; CAB56776.1; -  
DR PIR; B36364; B36364.  
DR PIR; B49647; B49647.  
DR HSP; P05713; 3RAB.  
DR Genew; HGNC:7007; MEL.  
DR MIM; 165040; -  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001230; Prenyl\_site.  
DR InterPro; IPR001806; Ras\_trnsfrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMS; TIGR00231; small\_GTP; 1.  
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;  
KW Proto-oncogene.  
FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).  
FT CONFLICT 177 183 LEGNSPQ -> WKATAP (IN REF. 2).  
SQ SEQUENCE 207 AA; 23668 MW; AA52DBF54A2CD056 CRC64;

Query Match 44.7%; Score 513.5; DB 1; Length 207;  
Best Local Similarity 50.0%; Pred. No. 2.9e-38;  
Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;  
Qy 25 SYDITGKVMLLDGTGVTCTFLQKDGAFSGTFTATVGNKVVVDGVRVKLQIW 84  
Db 4 TYDYLKLLIGDSGVGKTCVLFSEDAF NSTFISIGIDFKTITELDGKRIKLOIW 62  
Qy 85 DTAGQERFSTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEYQAQDVVIMLGNK 144  
Db 63 DTAGQERFTTAYRGAMGIMLYDITNEKSFNIRNIRNIBEHASADVEKMLGNK 122  
Qy 145 ADMSSERVISEDGETLAREYGVPELETSAKTMNVELAFIAIAKELYRAGHQAD---- 200  
Db 123 CDVNDKQVSKERGEKALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKLEGN 182  
Qy 201 EPSFOIRDYVESOKRSS 218  
Db 183 QGSNGVKITPDQQRSS 200

RESULT 8  
RAB1\_BETVU STANDARD; PRT; 215 AA.  
AC Q39433;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DE Ras-related protein RAB1BV.  
GN RAB1BV.  
OS Beta vulgaris (Sugar beet).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.  
OX NCBI\_TaxID=3555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. D100 KS 38080;  
RX MEDLINE=96307523; PubMed=8680960;

RA Dallery E., Cuief S., Ben Jilany K.E., Kerckaert J.-P., Hagege D.;  
RT "Molecular cloning and structural analysis of cDNAs that encode 3  
RT small GTP-binding proteins from sugar beet."  
RL C. R. Acad. Sci., III, Sci. Vie 319:91-97(1996).  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z49152; CAA89021.1; -  
DR HSP; P05713; 3RAB.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras\_trnsfrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMS; TIGR00231; small\_GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation.  
FT NP\_BIND 22 29 GTP (BY SIMILARITY).  
FT NP\_BIND 70 74 GTP (BY SIMILARITY).  
FT NP\_BIND 128 131 GTP (BY SIMILARITY).  
FT LIPID 212 212 GERANYL-GERANYL (BY SIMILARITY).  
FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).  
SQ SEQUENCE 215 AA; 23787 MW; 781CC5A17F29B8AA CRC64;

Query Match 44.6%; Score 512.5; DB 1; Length 215;  
Best Local Similarity 47.6%; Pred. No. 3.7e-38;  
Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;  
Qy 19 SPP--CSFSYDLTGKVMLLDGTGVTCTFLQKDGAFSGTFTATVGNKVVVDG 76  
Db 3 APPARADYDYLKLLIGDSGVGKSCLLRFSDGSETT-SFTITIGIDFKRIELDG 61  
Qy 77 VRVKLQIWDTAGQERFSTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEYQAQDV 136  
Db 62 KRIKLIQIWDTAGQERFRTTAYRGAMGIMLYDVTDESFNIRNIRNIEQHASDV 121  
Qy 137 VIMLGNKADM--SSERVISEDGETLAREYGVPELETSAKTMNVELAFIAIAKELK 195  
Db 122 NKILVGNKADMESKRAVPTAKQALADEYGIKFFETSAKTNLNVVEEVSFIARDIK 181  
Qy 196 G----HQADEFSQTRDYESOKK---RSSCC 220  
Db 182 ADSDTROEAQPSITIKPADQSGNQAAKSACC 213

RESULT 9  
RAB8\_DISOM STANDARD; PRT; 210 AA.  
ID RAB8\_DISOM  
AC P22128;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Ras-related protein Rab-8 (ORA2).  
OS Discopoge omnia (Electric ray).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Squala; Hypnosqualea; Pristioraja; Batoidea;  
OC Torpediniformes; Narcinoidei; Narcinidae; Discopogae.  
OX NCBI\_TaxID=7785;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Electric lobe;  
RX MEDLINE=91115900; PubMed=1899244;  
RA Ngsee J.K., Eiferink L.A., Scheller R.H.;  
RT "A family of ras-like GTP-binding proteins expressed in electromotor  
RT neurons."  
RL J. Biol. Chem. 266:2675-2680(1991).



RESULT 11	RESULT 12
ID RB8B_RAT	ID RB8B_HUMAN
AC P70550;	AC Q92930; Q9P293;
DT 01-NOV-1997 (Rel. 35, Created)	DT 15-JUL-1999 (Rel. 38, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)	DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)	DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-8B.	DE Ras-related protein Rab-8B.
GN RAB8B.	GN RAB8B.
OS Rattus norvegicus (Rat).	OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=10116;	OX NCBI_TaxID=9606;
RN [1]	RN [1]
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RM MEDLINE=96393028; PubMed=8799816;	RM MEDLINE=97182150; PubMed=9030196;
RA Armstrong J., Thompson N., Squire J.H., Smith J., Hayes B., Solari R.;	RA Chen D., Guo J., Gahl W.A.;
RT "Identification of a novel member of the Rab8 family from the rat	RT "Rab GTPases expressed in human melanoma cells.";
RT basophilic leukaemia cell line, RBL.2H3.";	RT Biochim. Biophys. Acta 1355:1-6(1997).
RL J. Cell Sci. 109:1265-1274(1996).	RL [2]
CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND	CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE (BY SIMILARITY).	CC NEUROTRANSMITTER RELEASE (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS OF EXPRESSION IN THE SPLEEN,	CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC TESTIS AND BRAIN.	CC [3]
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.	CC SEQUENCE OF 67-119 FROM N.A.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its	CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way	CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial	CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/	CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).	CC or send an email to license@isb-sib.ch).
CC EMBL; U53475; AAA99782.1; -	CC EMBL; AB038935; BAA92249.1; -
DR HSSP; P05713; 3RAB.	DR EMBL; BC020634; RAH20654.1; -
DR InterPro: IP0003579; GTPase_Rab.	DR HSSP; P05713; 3RAB.
DR InterPro: IP0001230; Prenyl_site.	DR InterPro: IP0003579; GTPase_Rab.
DR InterPro: IP0001806; Ras_trnsfrmg.	DR InterPro: IP0001230; Prenyl_site.
DR InterPro: IP0005225; Small_GTP.	DR InterPro: IP0001806; Ras_trnsfrmg.
DR Pfam; PF00071; ras; 1.	DR Pfam; PF0005225; Small_GTP.
DR PRINTS; PR00449; RASTRNSFRMNG.	DR Pfam; PF00071; ras; 1.
DR SMART; SM00175; RAB; 1.	DR PRINTS; PR00449; RASTRNSFRMNG.
DR TIGRFS; TIGR00231; small_GTP; 1.	DR SMART; SM00175; RAB; 1.
GW GTP-binding; Prenylation; Lipoprotein; Protein transport.	DR TIGRFS; TIGR00231; small_GTP; 1.
FT NP_BIND 15 22 GTP (BY SIMILARITY).	GW GTP-binding; Prenylation; Lipoprotein; Protein transport.
FT NP_BIND 63 67 GTP (BY SIMILARITY).	FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).	FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).	FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).	FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
SEQ SEQUENCE 207 AA; 23603 MW; 4A1A2B26BF9DCAF4 CRC64;	FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
Query Match 43.7%; Score 502.5; DB 1; Length 207;	SEQ SEQUENCE 207 AA; 23584 MW; 5960993C0F87F944 CRC64;
Best Local Similarity 47.3%; Pred. No. 2.7e-37;	Query Match 43.6%; Score 501; DB 1; Length 207;
Matches 97; Conservative 45; Mismatches 56; Indels 7; Gaps 3;	Best Local Similarity 49.2%; Pred. No. 3.7e-37;
QY 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVDGVRVKLIQW 84	Matches 98; Conservative 40; Mismatches 59; Indels 2; Gaps 2;
DB 4 TYDYLKLLIGDGSVGKTCFLFRSEDAF-NTTFISTIGIDFKRTTLDGKKIKLIQW 62	QY 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFSLGTFIATVGDIFRNKVVTVDGVRVKLIQW 84
QY 85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSFNIRAWLTFEIHAYQORDVIMLGNK 144	DB 4 TYDYLKLLIGDGSVGKTCFLFRSEDAF-NTTFISTIGIDFKRTTLDGKKIKLIQW 62
DB 63 DTAGQERFRITAYYRGAMGLMYDITNKSFNIRAWLTFEIHAYQORDVIMLGNK 122	QY 85 DTAGQERFRSVTHAYYRDAQALLLLYDITNKSFNIRAWLTFEIHAYQORDVIMLGNK 144
QY 145 ADMSSERVIRSEDEGTFLAREYGVPEFSAKTMGMVLEAFIAIAKELKYRHAQADEPSF 204	DB 63 DTAGQERFRITAYYRGAMGLMYDITNKSFNIRAWLTFEIHAYQORDVIMLGNK 122
DB 123 CDMDNRQVSKRGERGLADYGIKLETSKASTNVEEAFFTLARDIMTKLNRRKMDNS 182	QY 145 ADMSSERVIRSEDEGTFLAREYGVPEFSAKTMGMVLEAFIAIAKELKYRHAQADEPSF 204
QY 205 QIR---DIVESQKRRSC--CSFM 223	DB 123 CDMDNRQVSKRGERGLADYGIKLETSKASTNVEEAFFTLARDIMTKLNRRKMDNS 182
DB 183 SGAGGPVKITSRKSKTSFFRCSSL 207	QY 205 QIR---DIVESQKRRSC--CSFM 223

```
Db 63 DTAGQERFRITITAYRGAMGLVYDITNEKSFNKNWIRNIEBHASSDVERMILGNK 122
QY 145 ADMSSSRVRSDEGETLAREYGVFFLETSAKTMGNVLAFLAIKELKYRAGHQ-DEPS 203
Db 123 CDMDKRQVSKERGEKLAIDYGIKLETSKSSANVEAFFLARDIMTKLNKRNKDNS 182
QY 204 FQIRDYVESQKRSSCCSF 222
Db 183 AGAGGPVKITENRSKTSF 201

RESULT 13
RCL_ORYSA
ID RCL_ORYSA STANDARD; PRT; 202 AA.
AC P40392;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ras-related protein RCL1.
GN RCL1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Yamahoushi; TISSUE=Callus;
RX MEDLINE=94009718; PubMed=8405471;
RA Uchimiya H., Kidou S., Anal T., Umeda M., Aotsuka S., Tsuge T.,
RT Kato A.;
RA "Molecular structure of ras-related small GTP-binding protein genes
RT of rice plants and GTPase activities of gene products in Escherichia
RT coli.";
RL FEBS Lett. 332:282-286(1993).
CC -|- FUNCTION: POSSESSES GTPASE ACTIVITY.
CC -|- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S66160; AAB28535.1; -
DR HSP; S38740; S38740.
DR HSP; P01112; IPLK.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmng.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_GTP; 1.
DR TIGRFS; TIGR00231; small_GTP; 1.
KW NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 200 200 PALMITATE (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 202 AA; 22561 MW; 5A1CB5F93256B659 CRC64;

Query Match 43.5%; Score 500; DB 1; Length 202;
Best Local Similarity 46.3%; Pred. No. 4.4e-37;
Matches 94; Conservative 49; Mismatches 54; Indels 6; Gaps 3;

QY 23 SPYDITGKVMILGDTGVGKTCFLIOFKDGAFLSGFIATVGDIFRNKVVTVGVGRVKIQ 82
Db 2 NPEYDILFKLLITGDSGVGRSCLLRPADDSYLE-SYISTIGVDFKIRTVQDGKTIKIQ 60
```

```
QY 83 IWDTAGQERFRSVTHAYYRDAQAALLLLYDITNKSFNIRAWLTHIHEYAQRDVVIMLLG 142
Db 61 IWDTAGQERFRITTSYYRGAGHIIIVYDVTDQESFNNVKQWLNEIDRYASENVNKLVLG 120
QY 143 NKADSSSRVRSDEGETLAREYGVFFLETSAKTMGNVLAFLAIKELKYRAGHQ---- 198
Db 121 NKCDLAENRVVSYEAGKALADEIGIPFLETSAKDATNVEKAFMTWAGEIKNRMASQGRTN 180
QY 199 ADEPSFQIRDYVESQKRSSCCS 221
Db 181 ASKPA-TVQMPROPVQAQSSCCS 202

RESULT 14
SAS1_DICDI
ID SAS1_DICDI STANDARD; PRT; 208 AA.
AC P20790;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GTP-binding protein SAS1.
GN SAS1 OR SAS1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220623; PubMed=2109188;
RA Saxe S.A., Kimmel A.R.;
RT "SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum
RT with sequence similarities to essential genes in Saccharomyces
RT cerevisiae.";
RL Mol. Cell. Biol. 10:2367-2378(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89209367; PubMed=3149563;
RA Saxe S.A., Kimmel A.R.;
RT "Genes encoding novel GTP-binding proteins in Dictyostelium.";
RL Dev. Genet. 9:259-265(1988).
CC -|- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -|- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34457; AAA33249.1; -
DR PIR; A34716; A34716.
DR PIR; A61571; A61571.
DR HSP; P05713; 3RAB.
DR DictyDb; DD05020; sasA.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Palmitate.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT LIPID 207 207 PALMITATE (BY SIMILARITY).
FT LIPID 208 208 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 208 AA; 23243 MW; BF681EF5C810019E CRC64;

Query Match 43.3%; Score 498; DB 1; Length 208;
Best Local Similarity 46.6%; Pred. No. 6.8e-37;
```

```
Matches 95; Conservative 49; Mismatches 50; Indels 10; Gaps 3;
QY 23 SPSYDLTKVMLLGGTGVGKTCFLQKDGAFLSGTFIATVGIDFRNKVYTVDGVVRVQLQ 82
Db 9 SAAYDYLIKLLIGDSGVGKSCLLLRFSDFS-TPSFITIGIDFKIRTIIELEGGRIKIQ 67
QY 83 IWDTAGQERFRSRYTHAYYRDAQALLLLYDITNKKSSFDNRAMTEIHEYAQRDVVIMLIG 142
Db 68 IWDTAGQERFRITTYAYRGAGILLVYDVTDEKSFGRNMRIRNEQHATDSVNMKMLIG 127
QY 143 NKADMSSEIRVISEDGETLAREGVGFLETSAGTKGNVELAFLAIAKELKYRAGHQADSP 202
Db 128 NCDMAEKVVDSSRSKSLADETGKFLFLETSARNSINVEAFISLAKDIKKR---MIDTP 184
QY 203 SFQIR-----DYVESQKRSSCC 220
Db 185 NEQPQVVQPGTNGANNKKACC 208

RESULT 15
ID YPT2_MAIZE STANDARD; PRT; 203 AA.
AC Q05737;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE GTP-binding protein YPTM2.
GN YPTM2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Coleoptile;
RX MEDLINE=92115746; PubMed=1731354;
RA Palme K., Diefenthal T., Vington M., Sander C., Schell J.;
RT "Molecular cloning and structural analysis of genes from Zea mays
(L.) coding for members of the ras-related ypt gene family.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:787-791(1992).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: ITS EXPRESSION IS WEAK IN STEMS, HIGHER IN
CC ROOTS, LEAVES AND COLEOPTILES, BUT HIGHEST IN FLOWERS.
CC -!- PTM: THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
CC BIOLOGICAL FUNCTION.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X63278; CAA44919.1; .
CC PIR; B38202; B38202.
CC HSP; P05713; 3RAB.
CC MaizeDB; 78605; -.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001806; Ras_trnsfrmng.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; ras; 1.
CC PRINTS; PR00449; RASTRNSFRMG.
CC SMART; SM00175; RAB; 1.
CC TIGRfams; TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
```

```
FT LIPID 200 200 PALMITATE (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 203 AA; 22475 MW; E241326E7ACD1B8A CRC64;
Query Match 43.0%; Score 495; DB 1; Length 203;
Best Local Similarity 45.1%; Pred. No. 1.2e-36;
Matches 92; Conservative 51; Mismatches 53; Indels 8; Gaps 3;
QY 23 SPSYDLTKVMLLGGTGVGKTCFLQKDGAFLSGTFIATVGIDFRNKVYTVDGVVRVQLQ 82
Db 2 NPEYDYLFKLLIGDSGVGKSCLLLRFSDFS-YLSTIGVDFKIRTVQDQGTIKIQ 60
QY 83 IWDTAGQERFRSRYTHAYYRDAQALLLLYDITNKKSSFDNRAMTEIHEYAQRDVVIMLIG 142
Db 61 IWDTAGQERFRITTYAYRGAGILLVYDVTDEKSFGRNMRIRNEQHATDSVNMKMLIG 120
QY 143 NKADMSSEIRVISEDGETLAREGVGFLETSAGTKGNVELAFLAIAKELKYRAGHQ--- 198
Db 121 NKSDLTPANKVAVATETAKAFADENGIPFMETSAKNATNVQOAFMAAASIKDRMASQPAAA 180
QY 199 -ADEPSFQIRDYVESQKRSSCCS 221
Db 181 NARPATVQIRGQPVNOK--TSCCS 202

Search completed: November 17, 2002, 16:40:09
Job time : 26 sec's
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 17, 2002, 16:41:58 ; Search time 1955 Seconds  
(without alignments)  
1847.362 Million cell updates/sec

Title: US-09-817-199A-2  
Perfect score: 1150  
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKRSSCCSFM 223

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09817199/runat\_13112002\_134902\_5176/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817199.ecgn.1.1.899.ernat\_13112002\_134902\_5176 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1150	100.0	1014	14	BM921365	BM921365 AGENCOURT
2	1137	98.9	1051	12	EG283602	EG283602 602407930
3	1134	98.6	887	9	AL522282	AL522282 AL522282
4	1058	92.0	676	10	BB598938	BB598938 BB598938
5	1020	88.7	651	10	BB633978	BB633978 BB633978
6	1006	87.5	740	13	BI767046	BI767046 603054267
7	1001	87.0	655	10	BB635649	BB635649 BB635649
8	856.5	74.5	749	9	AL559085	AL559085 AL559085
9	856	74.4	733	10	BB619625	BB619625 BB619625
10	852	74.1	522	13	BM151643	BM151643 TCBAPIE10
c 11	829	72.1	736	13	EG934396	EG934396 SKI-0698
12	825	71.7	676	10	BB206788	BB206788 BB206788
13	790	68.7	553	12	BE755280	BE755280 209068 MA
14	763	66.3	488	13	BM149118	BM149118 TCARP2E63
15	760	66.1	888	14	EQ687186	EQ687186 AGENCOURT
16	760	66.1	953	12	BE906081	BE906081 601497118
17	754	65.6	895	14	BB688147	BB688147 AGENCOURT
18	752	65.4	579	10	BB621927	BB621927 BB621927
19	740	64.3	657	10	BB196489	BB196489 BB196489
20	739	64.3	455	10	AW464449	AW464449 BP230015B
21	728	63.3	691	12	BG253976	BG253976 602366910
22	721.5	62.7	672	13	BM492040	BM492040 p92n.pk0
23	650	56.5	701	10	AW914157	AW914157 EST345461
24	646	56.2	708	9	AJ442405	AJ442405 AJ442405
25	622	54.1	780	12	BG784310	BG784310 SEAMC004
26	618	53.7	443	10	AW956878	AW956878 EST368948
27	610	53.0	380	10	BE241772	BE241772 TCARP2E02
28	607	52.8	386	10	AW484587	AW484587 61493 MAR
29	588	51.1	820	13	BI916873	BI916873 603177832
30	582	50.6	782	13	BI545231	BI545231 603187353
31	581.5	50.6	658	10	AV697217	AV697217 AV697217
32	581	50.5	368	13	BM152052	BM152052 TCBAPIE12
33	580.5	50.5	648	10	AV697216	AV697216 AV697216
34	580.5	50.5	684	10	AV692701	AV692701 AV692701
35	580.5	50.5	697	10	AV692689	AV692689 AV692689
36	576.5	50.1	678	10	AV685389	AV685389 AV685389
37	572	49.7	668	10	AV688197	AV688197 AV688197
38	557.5	48.5	647	10	AV698293	AV698293 AV698293
39	557	48.4	678	10	AV686507	AV686507 AV686507
40	548	47.7	692	12	BG699266	BG699266 602678938
41	535	46.5	866	9	AU079251	AU079251 AU079251
42	530	46.1	592	13	BM023074	BM023074 id65h09.Y
43	528.5	46.0	767	12	BF608560	BF608560 MYL_00150
44	528.5	46.0	939	14	BQ947443	BQ947443 AGENCOURT
45	527	45.8	404	13	BM149006	BM149006 TCARP2E57

ALIGNMENTS

RESULT 1  
BM921365  
LOCUS  
DEFINITION AGENCOURT\_6626159 NIH\_MGC\_115 Homo sapiens linear EST 12-MAR-2002  
5', mRNA sequence.  
ACCESSION BM921365  
VERSION BM921365.1 GI:19371744  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1014)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12787 row: i column: 04  
High quality sequence stop: 735.  
Location/Qualifiers

## FEATURES

source

1. .1014  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5752779"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 69. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH\_MGC Library."

BASE COUNT 237 a 288 c 295 g 194 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,16e-139 Length: 1014  
Score: 1150.00 Matches: 223  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-817-199a-2 (1-223) x BM9211365 (1-1014)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
Db 48 ATACGGGACGCGGCGCGGTGGCCACCGGGATGGCGAGGCCCGCCGCGCTCCCG 107  
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
Db 108 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGAGACACAGCGCTC 167  
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
Db 168 GGCACCAACATGTTCTCTGATCCAAATCAAGAGCGGGCTTCTCTGTCGGAACCTTCATA 227  
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
Db 228 GCCACCGTGGCAGATACATTCAGGAACAAGGTGGTGCATCTGGATGGCGTGAGTGAAG 287  
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
Db 288 CTGCAGATCTGGACACCGCTGGCGAGGACCGGTTCCGAGGCGTACCCATGCTTATTAC 347  
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
Db 348 AGAGATGCTCAGGCCCTGCTCTGCTGTATGATACATCACCAACAATCTCTTTCGACAAC 407  
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140  
Db 408 ATCAGGCCCTGGCTCACTGAGATTATGATATGCGCCAGAGGACGCTGGTATCATGCTG 467  
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
Db 468 CTAGGCAACAAGCGGATATGACACGGAAGAGTATCCGTCCTCCGAACGAGGAGACC 527

QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
Db 528 TTGGCCAGGAGTACGGTGTTCCTTCTGGAGACAGCCCAAGACTGCATGATGTG 587  
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
Db 588 GAGTTAGCCTTTCTGGCCATCGCAAGAACTGAAATACCGGCGGCGCATCAGGCGGAT 647  
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
Db 648 GAGCCAGCTTCCAGATCCGAGACTATGTAGTCCAGAGAGGCGCTCCAGCTGCTGC 707  
QY 221 SerPheMet 223  
Db 708 TCCTTCATG 716

## RESULT 2

BG283602

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

TISSUE

CDNA

DNA

Clone

found

http://image.llnl.gov

Plate: LLAM10417 row: k column: 08

High quality sequence stop: 782.

FEATURES

source

1. .1051

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4520191"

/clone\_lib="NIH\_MGC\_91"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pCMV-SPORT6; Site.1: NotI;  
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.4 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 275 a 287 c 314 g 175 t

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3.24e-137

1137.00

99.55%

99.55%

98.87%

12

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

1051

223

0

1

0

US-09-817-199a-2 (1-223) x BG283602 (1-1051)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20

Db 3 ATGACGGGACGCGGCGCGGTGGCCACCGGGATGGCGAGGCCCGCCGCGCTCCCG 62



```
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
|||||
Db 63 CCTGCAGTCCGAGCTACAGCCTCAGGCGAAGGTGATGCTCTCGGAGACACAGCGCTC 122
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
|||||
Db 123 GGCACCAAGATGTTCTCTGATCCAAATCAAGACGCGGCTTCTCTCGGAACCTTCATA 182
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
|||||
Db 183 GCCACCGCTGGCATAGACTTCAGAACCAAGGTGGTACTGTGGATGGGTGAGATGAAG 242
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
|||||
Db 243 CTGCAGATCTGGCAGCACCGCTGGCAGGACGGTTCGGAAGCGTCACCCATCTATTAC 302
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
|||||
Db 303 AGAGATGCTCAGGCGCTTGTCTGTGTATGACATCACCAACAAATCTCTTTTCGACAC 362
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
|||||
Db 363 ATCAGGGCTGCTCAGTGAAGATTCATGATATGCCAGAGGAGCGTGGTATCATGCTG 422
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
|||||
Db 423 CTAGGCAACAAGCGGATAGACGCGAAGAGTGATCCGTTCCGAGACGCGAGACC 482
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
|||||
Db 483 TTGGCCAGGGAGTAGGGTGTCTCCCTTCTGGAGACGAGCGCAAGACTGGCATGAATGTG 542
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
|||||
Db 543 GAGTTAGCTTCTTGGCCATCCCAAGAACTGAATACCGGGCGGGCATCAGCGCGAT 602
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGln-LysLysArgSerSerCysCys 220
|||||
Db 603 GAGCCAGCTTCACATCCGAGACTATGTAGAGTCCAGAAAGACGCTCCAGCTGCTG 662
QY 220 sSerPheMet 223
|||||
Db 663 CTCCTTCATG 672

RESULT 3
AL522282 LOCUS AL522282 887 bp mRNA linear EST 13-FEB-2001
DEFINITION AL522282 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 5
prime, mRNA sequence.
ACCESSION AL522282
VERSION AL522282.1 GI:12785775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB008YF08"
/location_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
```

```
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 203 a 256 c 257 g 168 t
ORIGIN

Alignment Scores:
Pred. No.: 6 08e-137 Length: 887
Score: 1134.00 Matches: 220
Percent Similarity: 99.10% Conservative: 1
Best Local Similarity: 98.65% Mismatches: 2
Query Match: 98.61% Indels: 0
Db: 9 Gaps: 0

US-09-817-199a-2 (1-223) x AL522282 (1-887)
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
|||||
Db 28 ATGACGGGCGACGCCAGGCGCGCTTGCACCCGGGATGGCGAGCCGCCGAGCGCTCCCG 87
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
|||||
Db 88 CCTGCACTCCGAGCTACGACCTCACGGCAAGGTGATCTCTGGGAGACACAGCGCTC 147
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
|||||
Db 148 GCAAAACATGTTCTCTGATCCAAATCAAGACGGGCGCTTCTGTCCGGAACCTTCATA 207
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
|||||
Db 208 GCCACCGTCGGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCTGAAAGTGAAG 267
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
|||||
Db 268 CTGCAGATCTGGGACACCGCTGGGAGGAACGGTTCGGAAGCGTCAACCATGCTATTAC 327
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
|||||
Db 328 AGARATGCTCAGGCCCTTGGTCTGTGTATGACATCACCAACAAATCTCTTTCGACAC 387
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
|||||
Db 388 ATCAGGGCTGGCTCACTGAGATTCTATGATGATGCCAGAGGACGCTGGTATCATGCTG 447
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
|||||
Db 448 CTAGGCAACAAGCGGATATGAGCGGAARAGTATCCGTTCGGAACGAGAGACC 507
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
|||||
Db 508 TTGGCCAGGGAGTAGCGTGTCTCCCTTCCGAGAGACCGAGCCCAAGACTGGCATGAATGTG 567
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
|||||
Db 568 GAGTTAGCTTCTTGGCCATCCCAAGGAACCTGAATACCGGGCGGCGCATCAGGCGGAT 627
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
|||||
Db 628 GAGCCCAAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGC 687
QY 221 SerPheMet 223
|||||
Db 688 TCCTTCATG 696

RESULT 4
```

**BB598938** 676 bp mRNA linear EST 26-OCT-2001  
**LOCUS** BB598938 RIKEN full-length enriched, adult pancreas islet cells Mus musculus cDNA clone C820003E14 5', mRNA sequence.

**ACCESSION** BB598938.2 GI:16450366  
**VERSION** EST.  
**KEYWORDS** house mouse.  
**SOURCE** Mus musculus  
**ORGANISM** Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 676)

**REFERENCE** Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,Y., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.  
**AUTHORS** RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
**TITLE** Unpublished (2001)  
**JOURNAL** On Dec 1, 2000 this sequence version replaced gi:11507539.  
**COMMENT** Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaoka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata.K. and Hayashizaki,Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Location/Qualifiers  
1..676  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="C820003E14"  
/clone\_lib="RIKEN full-length enriched, adult pancreas islet cells"  
/tissue\_type="pancreas"  
/cell\_type="islet cells"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Site\_1: Sali; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia"

## RESULT 5

BB633978 651 bp mRNA linear EST 26-OCT-2001  
 LOCUS BB633978 RIKEN full-length enriched, adult male spinal cord Mus  
 DEFINITION musculus cDNA clone A330094K10 5', mRNA sequence.

ACCESSION BB633978  
 VERSION BB633978.1 GI:16470310

KEYWORDS EST.

SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

## AUTHORS

1 (bases 1 to 651)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,  
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,  
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
 Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

## TITLE

## JOURNAL

## COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9226

Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
 Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa,  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.

e mouse tissues.

## FEATURES

## Source

Location/Qualifiers  
 1...651  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="A330094K10"  
 /clone\_lib="RIKEN full-length enriched, adult male spinal  
 cord"  
 /sex="male"  
 /tissue\_type="spinal cord"  
 /dev\_stages="adult"  
 /lab\_host="DH10B"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN, Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 10.0 and subtraction to Rot = 459.0. Second  
 strand cDNA was prepared with the primer adapter of  
 sequence [5' GAGAGAGAGATCTCGATTAAATTAATATCCCTCCCCCCCC  
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLC I."

BASE COUNT 151 a 168 c 192 g 140 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,766-122 Length: 651  
 Score: 1020.00 Matches: 198  
 Percent Similarity: 95.75% Conservative: 5  
 Best Local Similarity: 93.40% Mismatches: 9  
 Query Match: 88.70% Indels: 0  
 Ds: 10 Gaps: 0

US-09-817-199A-2 (1-223) x BB633978 (1-651)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20

Db 16 ATGACTGGCACACACAGGAGCTGCTACCGCTGGGGATGGCAGGCCCTCAGCGCTCCCCG 75

Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40

Db 76 CCCTTCAGCCCCGAACCTACGATCTCACCGCAAGGTGATGCTCTTGGAGACTCGGGCGTC 135

Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60

Db 136 GGCACAAACCTGTTCTCTGATCCCAATTCAAAGACGGGGCTTCTGTCGCCGAACCTTCATA 195

Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80

Db 196 GCACCGCGGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGCCAGGGTGAAG 255

Qy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100

Db 256 CTTTCAGATCTGGGACACTCAGACAGGAGCGCTCCGCGAGTGTACCCATGCTTATTAC 315

Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120

Db 316 CGAGATGCTCAGGCTTTGCTCTGTGTATGATCATCACCAACAGTCTCTTTTGACAAC 375

Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140

Db 376 ATCAGGGCTGGCTCACAGAGATTTCATGAGTATGCCACAGAGACGTGGTGATTGCTT 435

Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160

Db 436 CTAGGCAACAAGGCCGATTAAGACGCGAAAGGGTGTATCCGTTCTGAAGATGGAGAGAC 495

Qy 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180

Db 496 CTGGCCAGGGAATATGGTGTCTCTTCATGGAGACAGTGGCCAAAGACTGGCATGAACGTG 555

Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200

Db 556 GAGTGGCCCTTCTGGGCATTGCCAAGGAACCTGAATACCGTGCAGGGAGGAGCGCTGAT 615

Qy 201 GluProSerPheGlnIleArgAspTyrValGluSer 212

Db 616 GAGCCAGCTTTCAGATCCGAGACTATGTGGAGTCC 651

RESULT 6

BI767046

LOCUS

DEFINITION 603054267F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5203600 5',

linear EST 25-SEP-2001

mRNA 740 bp



Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alzawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001). Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

#### FEATURES

Location/Qualifiers  
 1. .655  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="M430077P13"  
 /clone\_lib="RIKEN full-length enriched, 0 day neonate thymus"  
 /tissue\_type="thymus"  
 /dev\_stage="0 day neonate"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot -10.0 and subtraction to Rot -459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCCAGTAAATTAATATATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 152 a 171 c 191 g 141 t

ORIGIN

Alignment Scores:  
 Pred. No.: 8.4e-120 Length: 655  
 Score: 1001.00 Matches: 194  
 Percent Similarity: 94.34% Conservative: 6  
 Best Local Similarity: 91.51% Mismatches: 12  
 Query Match: 87.04% Indels: 0  
 DB: 10 Gaps: 0

US-09-817-199A-2 (1-223) x BB635649 (1-655)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
 Db 18 ATGACTGGCACACAGGAGGCTCTACCGCTGGGGATGGCGAGGCCCTTCGAGCGCTCCCG 77  
 Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
 Db 78 CCCTTCACCGCGAAGTACAGTCTACCGCGAAGGTGATGCTCTTGGAGACTCGGGCGTC 137  
 Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
 Db 138 GGCMAAACCTGTTCTCTGATCCAAATCAAGAGCGGGCTCTCTGTCGGAACCTTCATA 197  
 Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
 Db 198 GCCACCGTCGCATAGACTTCAGGAATAAGTGTGACAGTGGATGGTCCAGGGTGAAG 257  
 Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrThr 100  
 Db 258 CTTAGATCTGGACACTGCAGGACAGGAGCGCTTCCCGAGGTGACCCATGCTTATTAC 317  
 Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120

Db 318 CGAGATGCTCAGCGCTTCTCTCTGTGTATGACATCACCAACCGACTCTCTTTTGACAAC 377  
 Qy 121 IleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
 Db 378 ATCAGGCGCTGGCTCACAGAGATTCATGATGATGCCAGAGAGACGTGGTGATTATGCTT 437  
 Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
 Db 438 CTAGGCAACAAGCGCGATGTAAGCAGCGCAAGGGGTGATCCGTTCTGAAGATGGAGACGA 497  
 Qy 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
 Db 498 CTGGCCAGGAATATGTTCTCTCTTCATGGAGACGAGTCCGCAAGACTGGCATGAACGTG 557  
 Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
 Db 558 GAGTTGGCTTCTCTGGGCAATGCCCAAGAACTGAATACCGTGCAGGAGGCGCTGAT 617  
 Qy 201 GluProSerPheGlnIleArgAspTyrValGluSer 212  
 Db 618 GAGCCCCAGCTTCAGATCCGAGACTATGTGGAGTCC 653

RESULT 8  
 AL559085 749 bp mRNA linear EST 16-FEB-2001  
 LOCUS AL559085 LTI\_NFL008\_TC2 Homo sapiens cDNA clone CS0DJ014YB11 5  
 DEFINITION prime: mRNA sequence.  
 ACCESSION AL559085  
 VERSION AL559085.1 GI:12904236  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (Bases 1 to 749)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).  
 Location/Qualifiers  
 1. .749  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DJ014YB11"  
 /clone\_lib="LTI\_NFL008\_TC2"  
 /sex="male"  
 /tissue\_type="T cells from T cell leukemia"  
 /note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT 157 a 213 c 228 g 129 t 22 others

ORIGIN

Alignment Scores:  
 Pred. No.: 7.39e-101 Length: 749  
 Score: 856.50 Matches: 180  
 Percent Similarity: 79.22% Conservative: 3  
 Best Local Similarity: 77.92% Mismatches: 36  
 Query Match: 74.48% Indels: 12  
 DB: 9 Gaps: 1

US-09-817-199A-2 (1-223) x AL559085 (1-749)



```
Db 3 ACTGCACACTAGGAGCTGCTACCCCTGGGGATGCGGAGGCCCTTCAGCGCTCCCGCC 62
Qy 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 63 TTCACCCGCACTAGCATCTCACCGCAAGGTGATGCTCTTGGAGACTCGCGCGCTGGC 122
Qy 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db 123 AAACCTGCTTCTCATCAATTCAAAGACGGGGCTCTCTCGCGGAACCTTCATAGCC 182
Qy 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 183 ACCCTGCGCATAGACTTCAGGAATAAGTGGTGACATGGATGGTCCCGAGGTGAAGCTT 242
Qy 82 GlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 243 CAGATCTGGGACATGCGAGGACAGAGCGCTTCCCGAGTGTGACCCATGCTTTATCCGA 302
Qy 102 AspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 303 GATGCTCAAGCTTTGCTCTGTGTGATGACATCACCAACCAATCTTTTTCACAACATA 362
Qy 122 ArgAlaTyrLeuThr-GluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 141
Db 363 AGGGCTGGCTTACAAAAGATTTCATGAAGTTCGCCAGAGAGACGTGGTGATTATGCTTTT 422
Qy 141 uGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrIle 161
Db 423 AGGCAC-AAGGCGGATGTAAGCAGCGAAAGGGTGATCGCTTCTGAAGAGGAACACT 481
Qy 161 uAlaArgGluTyrGlyValProPheLeuGlu-ThrSerAlaLysThrGlyMetAsnValG 181
Db 482 GGCCAGGAATATGTTGGTCTTCATGAGAACCGTGGCCAAAGACTGGCTTGAACGTGG 541
Qy 181 uLeuAla-PheLeuAlaIleAlaLysGlu-LeuLysTyrArgAla-GlyHisGlnAlaA 200
Db 542 AGTTGGCCCTTCTGGCAATTGCCAAGGAACTTAAATACCGTGAAGGAGCGACTCTG 601
Qy 200 spGluProSerPheGlnIleArgAspTyrValGlu-SerGlnLysLysArgSerSerCys 219
Db 602 ATGACACCCACTTTCAAAACCGAGATATGTGGAGTCCCAAAAACCGCTCCCAACTGG 561
Qy 220 Cys 220
Db 662 TGC 664

RESULT 10
BM151643
LOCUS
DEFINITION
522 bp mRNA linear EST 30-NOV-2001
TCBAP1E10302 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP1030, mRNA
sequence.
ACCESSION
BM151643
VERSION
BM151643.1 GI:17175247
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 522)
AUTHORS
Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
Gunarathne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
TITLE
Pediatric Leukemia cDNA Sequencing Project (2001)
JOURNAL
Unpublished (2001)
COMMENT
Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Seq primer: M13 primer.
Location/Qualifiers
```

```
source
1..522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCBAP1030"
/clone_lib="pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project-TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/Note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGAGCTCGGATCCGCCGCCCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P. Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasakini, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT 119 a 146 c 158 g 99 t
ORIGIN
Alignment Scores:
Pred. No.: 1.6e-100 Length: 522
Score: 852.00 Matches: 165
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.09% Indels: 0
DB: 13 Gaps: 0
US-09-817-199A-2 (1-223) x BM151643 (1-522)
Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 19 ATGAGCGGCGACCGAGCGCGCTTGCCACCGGGATGGGAGCGCGCTCCCG 78
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 79 CCCTGCAGTCCGAGCTACGACCTCAGGGGCAAGGTGATGCTTCTGGAGACACACAGCGTC 138
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 139 GGCAAAACATGTTCTGTATCCAAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCATA 198
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 199 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 258
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 259 CTGCAGATCTGGACACCCCTGGGAGGAACGGTTCGGAAGCGTCAACCATGCTTTATTAC 318
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 319 AGAGATGCTCAGCGCCCTGCTTCTGCTGTATGACATCACCAACAAATCTTCTTTCGACAAC 378
Qy 121 IleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 379 ATCAGGGCTGGCTCACTGAGATTATGATGATGCCAGAGGACGTGGTGATCATGCTG 438
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 439 CTAGGCCAACAGCGGATATGAGCAGCGGAAGAGTATCGTTCGCAAGACGAGAGACC 498
Qy 161 LeuAlaArgGluTyr 165
|||||
```

```

Db 499 TTGCCAGGAGTAC 513
RESULT 11
BG934396/c
LOCUS
DEFINITION
  BG934396 736 bp mRNA linear EST 02-OCT-2001
  SK1-0698 Atlantic Salmon kidney SalmO salar cDNA clone SK1-0698
  similar to GTPase Rab37, mRNA sequence.
ACCESSION
  BG934396
VERSION
  BG934396.1 GI:15844224
KEYWORDS
  EST.
SOURCE
  Atlantic salmon.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
  Protacanthopterygii; Salmoniformes; Salmonidae; SalmO.
REFERENCE
  1 (bases 1 to 736)
  Douglas, S.E., Tsol, S.C.M., Penny, S., Melville, K., Reith, M.E. and
  Ewart, K.V.
  Expressed Sequence Tags - A Snapshot of the Fish Genome
  Unpublished (2001)
  Contact: Douglas S
  Genome Sciences
  NRC Institute for Marine Biosciences
  1411 Oxford St., Halifax, Nova Scotia B3H3Z1 Canada
  Tel: (902) 426-4991
  Fax: (902) 426-9413
  Email: susan.douglas@nrc.ca
  Seq primer: SK.
FEATURES
  source
    1..736
    /organism="Salmo salar"
    /db_xref="taxon:8030"
    /clone="SK1-0698"
    /clone_lib="Atlantic Salmon kidney"
    /dev_stage="adult"
    /note="Organ: kidney; Vector: Lambda ZAP II"
BASE COUNT
  146 a 220 c 180 g 190 t
ORIGIN
  Alignment Scores:
  Pred. No.: 2,78e-97 Length: 736
  Score: 829.00 Matches: 155
  Percent Similarity: 90.26% Conservative: 21
  Best Local Similarity: 79.49% Mismatches: 19
  Query Match: 72.09% Indels: 0
  DB: 13 Gaps: 0
US-09-817-199A-2 (1-223) x BG934396 (1-736)
QY 26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
  ::::: |||||
Db 666 TTGTATTCCTTCAGGTGATGCTGCTGGGAGACTCTGCGGTGGGGAGACATGTGTG 607
QY 46 LeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIle 65
  ::::: |||||
Db 606 CTGGTCGCTTTAAAGACGGGCGCTTTTGGGAGGCACTTCATAGCCACCGTGGGAATA 547
  ::::: |||||
QY 66 AspPheArgAsnLysValThrValAspGlyValArgValLysLeuGlnIleTrpAsp 85
  ::::: |||||
Db 546 GACTTTAGGAATAAAGTTGTGCGGTGGACAAACGGAGGTCAAACTCCAGATCTGGGAT 487
  ::::: |||||
QY 86 ThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAla 105
  ::::: |||||
Db 486 ACTGCAGGACAGAGAGATTCGCGAGTGCACGCGCTACTACAGACGCACAGGCC 427
  ::::: |||||
QY 106 LeuLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsnIleArgAlaTrpLeu 125
  ::::: |||||
Db 426 CTCCTCCTGCTATGACATCACCAGCAAGTCATCTTTTGACACATCATGAGCCCTGGCTG 367
  ::::: |||||
QY 126 ThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuGlyAsnLysAla 145
  ::::: |||||
Db 366 ACTGAATCCATCAGTATGCTCAGAGGATGTGGTCATCATGTGTGTCGCAACAAGACG 307
  ::::: |||||

```

```

QY 146 AspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyr 165
  ::::: |||||
Db 306 GACATGCCGAGAGAGGTCATCAAGCGTGAGAGGAGAGAGCTGGCCAGGATAT 247
  ::::: |||||
QY 166 GlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeu 185
  ::::: |||||
Db 246 GGAGTCCCTTTCATGTGAGACCACTGGCAAGCTGAGTCAACGTAGAGCTGGCTCTCTG 187
  ::::: |||||
QY 186 AlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGln 205
  ::::: |||||
Db 186 GCTGTAGCAAGAGTTGAAGCAGACAGCTGCCAGACCCAGCCCAAGTTCACG 127
  ::::: |||||
QY 206 IleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
  ::::: |||||
Db 126 ATCCAGCACTACATCGCTGTGAGAGCAGCAAGCTCTGGCTGTGT 82
  ::::: |||||
RESULT 12
BB206788
LOCUS
DEFINITION
  BB206788 RIKEN full-length enriched, 0 day neonate thymus Mus
  musculus cDNA clone A430077P13 3', mRNA sequence.
ACCESSION
  BB206788
VERSION
  BB206788.2 GI:16352175
KEYWORDS
  EST.
SOURCE
  house mouse.
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 676)
  Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
  Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
  M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
  Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
  D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
  Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, F., Toya, T.,
  Muramatsu, M. and Hayashizaki, Y.
  RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
  Unpublished (2001)
  On Jun 30, 2000 this sequence version replaced gi:8871741.
  Contact: Yoshihide Hayashizaki
  Laboratory for Genome Exploration Research Group, RIKEN Genomic
  Sciences Center (GSC), Yokohama Institute
  The Institute of Physical and Chemical Research (RIKEN)
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  Tel: 81-45-503-9222
  Fax: 81-45-503-9216
  Email: genome-res@gsc.riken.go.jp
  URL: http://genome.gsc.riken.go.jp/
  Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
  M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new
  genes. Genome Res. 10 (10), 1617-1630 (2000)
  wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
  Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura,
  S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
  Hayashizaki, Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multipipillary sequencer. Genome Res.
  10 (11), 1757-1771 (2000)
  Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
  Y. and Hayashizaki, Y.
  Computer-based methods for the mouse full-length cDNA
  nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
  Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alizawa,
  K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
  Hayashizaki, Y.
  Computational Analysis of Full-Length Mouse cDNAs Compared with
  Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
  Please visit our web site (http://genome.gsc.riken.go.jp/) for
  further details.
  cDNA library was prepared and sequenced in Mouse Genome

```



Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

## FEATURES

source

Location/Qualifiers

1..676.  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="A430077P13"  
/clone\_lib="RIKEN full-length enriched, 0 day neonate  
thymus"  
/tissue\_type="thymus"  
/dev\_stage="0 day neonate"  
/lab\_host="DH10B"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3']. cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 10.0 and subtraction to Rot = 459.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGATCTCGAGTTAATTAATATCCCCCCCCC  
3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
modified pBluescript KS(+) after bulk excision from Lambda  
FLC I."

BASE COUNT 166 a 167 c 196 g 147 t

## ORIGIN

## Alignment Scores:

Pred. No.: 8,07e-97 Length: 676  
Score: 825.00 Matches: 161  
Percent Similarity: 98.21% Conservative: 4  
Best Local Similarity: 95.83% Mismatches: 3  
Query Match: 71.74% Indels: 0  
DB: 10 Gaps: 0

US-09-817-199a-2 (1-223) x BB206788 (1-676)

Qy 56 SerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAsp 75  
|||||  
Db 3 TCGGGNACCTTCATAGCCACCGTCGGCATGACTTCAGGAATAAGTGTGACAGTGGAT 62  
|||  
Qy 76 GlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerVal 95  
|||  
Db 63 GGTGCCAGGGTGAAGCTTCAGATCTGGGACACTGCAGGACAGAGCGCTTCGCAGTGG 122  
|||  
Qy 96 ThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLys 115  
|||||  
Db 123 ACCCATGCTTATACCGAGATGCTCAGGCTTTCCTCTTGTATGACATCACCACAG 182  
|||  
Qy 116 SerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAsp 135  
|||||  
Db 183 TCTCTCTTTTGACAAATCATCAGGCGCTGGCTCAGAGATTCATGAGTATGCCAGAGAGAC 242  
|||  
Qy 136 ValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSer 155  
|||||  
Db 243 GTGGTGATTTGCTTCTAGCAACAGCGCGATGAAGCAGCGAAGGGTGATCCGTTCT 302  
|||  
Qy 156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLys 175  
|||||  
Db 303 GAAGATGGAGACACTGGCCAGGAATATGCTTCTTCATGAGACACAGTGCCTCAAG 362  
|||  
Qy 176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195  
|||||  
Db 363 ACTGGCATGAACGTGGAGTGGCCCTTCCTGGCAATTTGCCAAGAACTGAAATACCGTGA 422  
|||  
Qy 196 GlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLys 215  
|||||

Db 423 GGGAGGACGCTGTAGTGGAGCCAGCTTCCAGATCCGAGACTATGTGGATCCCAAGAAG 482  
||| |||  
Qy 216 ArgSerSerCysCysSerPheMet 223  
|||||  
Db 483 CGCTCCAGCTGCTGCTCTTTGTG 506  
|||||  
RESULT 13  
BE755280 553 bp mRNA linear EST 25-APR-2001  
LOCUS 209068 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
DEFINITION BE755280  
ACCESSION BE755280  
VERSION BE755280.1 GI:10169272  
KEYWORDS EST.  
SOURCE COW.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 553)  
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett,  
G.D., Heaton,W.P., Laegreid,W., Rohrer,G.A., Chitko-Mckown,C.G.,  
Perteau,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and  
Keele,J.W.  
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
JOURNAL Genome Res. 11 (4), 626-630 (2001)  
MEDLINE 21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred 18  
and v0.980904.e. Vector identified by cross\_match with the -minscore 15  
and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCACGACG  
Plate: 57 row: E column: 12  
Seq primer: ATTTAGTGACACTATAG.  
Location/Qualifiers  
FEATURES  
source  
1..553  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, thymus,  
semitendinosus muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."  
BASE COUNT 111 a 167 c 174 g 101 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.116e-92 Length: 553  
Score: 790.00 Matches: 151  
Percent Similarity: 97.52% Conservative: 6  
Best Local Similarity: 93.79% Mismatches: 4  
Query Match: 68.70% Indels: 0  
DB: 12 Gaps: 0  
US-09-817-199a-2 (1-223) x BE755280 (1-553)  
Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
|||||  
Db 71 ATACCGGAAACGCTGGACCGGCCACCGGGATGGGAGGCCGCCCGAGCGCTCCCA 130  
|||||  
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
|||||

Db 131 CCCTGGGGCGGACCGATCTCACGGGCAAGGTGATGCTTCTGGGAGACTCGGGGCTC 190

QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
|||||.....:::|||||.....

Db 191 GGCAAAACCTGTTTCTTCATCCGATTCAGAGACGGGCGCTTCTGTCGGGACCTTCATA 250

QY 61 AlaThrValGlyIleAspPheArgAsnLysValThrValAspGlyValArgValLys 80  
|||||.....

Db 251 GCCACCGTCGGCATAGACTTCAGAGCAAAAGTGGTACCGTGGATGGTGAGAGTGA 310

QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
|||||.....

Db 311 CTCAGATCTGGGACACGGCAGGCGGTTCGCGAGTGTGCCCATGCTTATTAC 370

QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
|||||.....

Db 371 CGAGACCCCGAGCGCTTGCCTGCTGTACGATCACCACAAATCTTCTTTGACAAC 430

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
|||||.....

Db 431 ATCCGGCTTGGCTCACTGAGATTCAGAGTACCGCCAGAGGATGGTGTATCATGCTT 490

QY 141 LeuGlyAsnLysAlaAspMetSerGluArgValIleArgSerGluAspGlyGluThr 160  
|||||.....

Db 491 CTGGCAACAAGCAGATGTGACACTGAAGGGTGTATCGCTCAGAGCAGGAGATG 550

QY 161 Leu 161  
|||

Db 551 CTG 553

RESULT 14

LOCUS BM149118

DEFINITION TCAAP2E6371 Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP6371, mRNA sequence.

ACCESSION BM149118

VERSION BM149118.1 GI:17170448

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr., Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.

TITLE Pediatric Leukemia cDNA Sequencing Project (2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Dr. Judith F. Margolin  
Texas Children's Cancer Center and Human Genome Sequencing Center  
at Baylor College of Medicine  
1102 Bates, MC3-3320 Houston, TX 77030, USA  
Tel: 832-824-4536  
Fax: 832-825-4038  
Email: clones@xccc.org  
Seq primer: M13 primer.  
Location/Qualifiers  
1.488  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="TCAAP6371"  
/clone\_lib="Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA"  
/sex="male"  
/tissue\_type="leukopheresis"  
/cell\_type="myeloid cell"  
/dev\_stage="pediatric 6 years"  
/lab\_host="DH10B"  
/note="Vector: lambda pSB; Site\_1: BamHI; Site\_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer [5'-GGAGGACTCAGCGCCGAGGAGGAG(T)VN 3'; V-A,C,G; N-A,C,G,T] and then dg tailed. Second strand was primed with a BamHI-dC primer

[5'AGAGAGCTCGGATCCGGCGCCGCAATAATAAT(C)3'].  
Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and SalI sites of lambda pSB vector. Library was constructed by Wei Yu at RIKEN normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T, Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y, High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997).

BASE COUNT 112 a 134 c 145 g 96 t  
ORIGIN

## Alignment Scores:

5.88e-89 Length: 488  
Pred. No.: 763.00 Matches: 152  
Score: 763.00  
Percent Similarity: 97.48% Conservatives: 3  
Best Local Similarity: 95.60% Mismatches: 4  
Query Match: 66.35% Indels: 1  
DB: 13 Gaps: 0

US-09-817-199a-2 (1-223) x BM149118 (1-488)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
|||||.....

Db 12 ATGACGGGACCGCCAGCGCGCTTGCACCGGGATGGCGAGGCCCGGCGCTCCCG 71  
|||||.....

QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuGlyAspThrGlyVal 40  
|||||.....

Db 72 CCTCGCAGTCCGAGCTTTT-CTCACGGCAAGGTGATGCTTCTGGGACACAGCGCTC 130  
|||||.....

QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
|||||.....

Db 131 GGCAAAACATGTTTCTGATCAATTCAAAGACGGCGCTTCTGTCGGGACCTTCATA 190  
|||||.....

QY 61 AlaThrValGlyIleAspPheArgAsnLysValThrValAspGlyValArgValLys 80  
|||||.....

Db 191 GCCACCGTCGGCATAGACTTCAGGACAAAGTGGTGTGATGGCTGGAGTGAAG 250  
|||||.....

QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
|||||.....

Db 251 CTCAGATCTGGGACACCGCTGNGCAGGACCGTTCGAGCGTCAACCATGCTTATTAC 310  
|||||.....

QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
|||||.....

Db 311 AGAGATGCTCAGCGCTTGTCTGTATGATCATCACCAACAACTCTTCTTCGACAAC 370  
|||||.....

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
|||||.....

Db 371 ATCAGGCGCTGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 430  
|||||.....

QY 141 LeuGlyAsnLysAlaAspMetSerGluArgValIleArgSerGluAspGlyGlu 159  
|||||.....

Db 431 CTAGCAACAGGGCGGATATGAGCAGCGAAGAGTATCCGTTCCGAGACCGAGAG 487  
|||||.....

## RESULT 15

BQ687186

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

## FEATURES

source

BQ687186 888 bp mRNA linear EST 15-JUL-2002  
AGENCOURT\_8034867 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6206869  
5', mRNA sequence.  
BQ687186  
BQ687186.1 GI:21812502  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 888)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM2362 row: a column: 14  
 High quality sequence stop: 662.

FEATURES  
 source  
 1..888  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6206869"  
 /clone\_lib="NIH\_MGC.110"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGCAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

BASE COUNT 172 a 279 c 289 g 148 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.77e-88 Length: 888  
 Score: 760.00 Matches: 149  
 Percent Similarity: 76.62% Conservative: 28  
 Best Local Similarity: 64.50% Mismatches: 30  
 Query Match: 66.09% Indels: 24  
 Db: 14 Gaps: 3

US-09-817-199A-2 (1-223) x B0687186 (1-888)

Qy 4 ThrProGlyAla-----ValAlaThrArgAspGlyGluAlaProGluArgSer----- 19  
 Db 81 ACCCCCGCTGCCTCCACGCTGCCACCGCCACCGGGCCGCGCGGCTCCGGGACT 140  
 Qy 20 -----ProProCysSerPro----- 24  
 Db 141 GCGCTTTCGGCCGCGCGCGCCCAACGGCCCTTCACGCGCGCGCGCTCGCTT 200  
 Qy 25 -----SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39  
 Db 201 GCGCGCGGTGTCGACTTCTACGACGTCGCTTCAAGGTCATGCTGGTGGGACTCGGT 260  
 Qy 40 ValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59  
 Db 261 GTGGGAGACCTGTCTGTGTGGATTCAGGATGGTCTTCTGCGGGGACCTTC 320  
 Qy 60 IleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgVal 79  
 Db 321 ATCTCCACCGTAGGCATTGACTTCCGGAAACAAAGTTCTGACGTGGTGTGAAGGTG 380  
 Qy 80 LysLeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 99  
 Db 381 AAGCTGCAGATGTGGGACACAGCTGCTCAGGAGCGGTTCCCGAGTGTACCCATGCCCTAC 440  
 Qy 100 TyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheasp 119  
 Db 441 TACCGGGATGCTCATGCTCTGCTGTCTACGATGTACCAACAAAGGCTCCTTTGAC 500  
 Qy 120 AsnIleArgAlaThrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMet 139  
 Db 501 AACATCCAGGCTTGCTGACCGAGATCCACGAGTACGCCACACGACGTCGCGCTCATG 560  
 Qy 140 LeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159  
 Db 561 CTGCTGGGGAACAAGGTGGACTCTCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGAG 620

Qy 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179  
 Db 621 AAGCTGGCCCAAGGAGTATGGACTGCCCTTCATGGAGACCGCCCAAGACGGGCTCAAC 680  
 Qy 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla 199  
 Db 681 GTGGACTTGGCTTTCACAGCCATAGCAAGAGGAGTGAAGCAGCGCTCCCTGAAGGCTCCC 740  
 Qy 200 AspGluProSerPheGlnIleArgAspTyrVal 210  
 Db 741 ACGAGGCGCGCTTCCGGCTGCATGATTACGTT 773

Search completed: November 17, 2002, 18:05:43  
 Job time : 1959 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 16:01:33 ; Search time 81 Seconds  
(without alignments)  
567.266 Million cell updates/sec

Title: US-09-817-199A-2  
Perfect score: 1150  
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSCGSEFM 223

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertibrate.\*
- 14: sp.unclassified.\*
- 15: sp.rvirus.\*
- 16: sp.bacteriap.\*
- 17: sp.archheap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	697	60.6	666	5	Q9VP48	Q9VP48 drosophila
2	648.5	56.4	388	5	Q35R32	Q95r32 drosophila
3	533	46.3	196	5	O02046	O02046 caenorhabdi
4	521.5	45.3	207	11	O8VCF6	O8vcf6 mus musculus
5	514.5	44.7	202	10	Q9SXT5	Q9sxt5 cicier arlet
6	513.5	44.7	202	10	Q08153	Q08153 pisum sativ
7	510	44.3	216	10	O49844	O49844 daucus caro
8	509.5	44.3	214	10	Q40218	Q40218 lotus japon
9	509	44.3	203	10	Q94952	Q94952 oryza sativ
10	509	44.3	216	10	Q40215	Q40215 lotus japon
11	509	44.3	218	10	Q9SF91	Q9sf91 arabidopsis
12	508.5	44.2	204	5	O15971	O15971 drosophila
13	507.5	44.1	202	10	Q40203	Q40203 lotus japon
14	507	44.1	216	10	Q9LZD4	Q9lzd4 arabidopsis
15	507	44.1	216	10	O8VWF9	O8vwf9 nicotiana t
16	506	44.0	203	10	Q9M7P5	Q9m7p5 capsicum an

17	506	44.0	216	10	Q40177	Q40177 lycopersico
18	505.5	44.0	201	5	Q94148	Q94148 caenorhabdi
19	505	43.9	203	10	O24112	O24112 nicotiana p
20	505	43.9	216	10	O8W3J4	O8w3j4 nicotiana t
21	504	43.8	216	10	O8W3J3	O8w3j3 nicotiana t
22	503	43.7	203	10	Q41340	Q41340 lycopersico
23	502	43.7	216	10	O9FJF1	O9fjfl arabidopsis
24	501.5	43.6	203	10	O8RU63	O8ru63 oryza sativ
25	501	43.6	203	10	O940Z7	O940z7 arabidopsis
26	500	43.5	216	10	O24466	O24466 arabidopsis
27	499.5	43.4	216	10	O8W3J2	O8w3j2 nicotiana t
28	499	43.4	212	10	Q40217	Q40217 lotus japon
29	498	43.3	203	10	Q41338	Q41338 lycopersico
30	498	43.3	216	10	O9SWV8	O9swv8 arabidopsis
31	497.5	43.3	202	10	O9FPJ4	O9fpj4 arabidopsis
32	497.5	43.3	202	10	O08155	O08155 pisum sativ
33	497.5	43.3	215	10	Q41023	Q41023 pisum sativ
34	496.5	43.2	202	10	O9SEH3	O9seh3 arabidopsis
35	495	43.0	215	10	Q40219	Q40219 lotus japon
36	495	43.0	216	10	Q41024	Q41024 pisum sativ
37	490.5	42.7	215	10	Q41022	Q41022 pisum sativ
38	489	42.5	222	10	O96362	O96362 brassica ca
39	488	42.4	203	10	O08154	O08154 pisum sativ
40	488	42.4	203	10	Q40569	Q40569 nicotiana t
41	487.5	42.4	202	10	O8W4S8	O8w4s8 arabidopsis
42	487.5	42.4	207	5	O18338	O18338 drosophila
43	487	42.3	200	4	Q9H0T3	Q9h0t3 homo sapien
44	486	42.3	202	10	Q9ZRH6	Q9zrh6 petunia hyb
45	484.5	42.1	215	10	Q41061	Q41061 pisum sativ

ALIGNMENTS

RESULT 1

Q9VP48	PRELIMINARY;	PRT;	666 AA.
ID	Q9VP48;		
AC	Q9VP48;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	CG7605 protein.		
GN	CG7605.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY.		
RX	MEDLINE=20190006; PubMed=10731132;		
RA	Adams M.D., Celisniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abri J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Fosler K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		



[illegible]

RT	"Isolation of a cDNA encoding a rab-related small GTP-binding protein.					
RT	(Accession No. AJ001367) from carrot cells (PGR97-185)."					
RL	Submitted (OCT-1997) to the EMBL/GenBank/DDJB databases.					
CC	-1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.					
DR	EMBL; AJ001367; CAA04701.1; -;					
DR	HSP; P05713; 3RAB.					
DR	InterPro: IPR003579; GTPase_Rab.					
DR	InterPro: IPR001806; Ras.transfmg.					
DR	InterPro: IPR002078; Sig54.interact.					
DR	InterPro: IPR005225; Small_GTP.					
DR	Pfam: PF00071; ras; 1.					
DR	PRINTS; PR00449; RASTRNSFRMG.					
DR	SMART; SM00175; RAB; 1.					
DR	TIGRFAMS; TIGR00231; small_GTP; 1.					
DR	PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.					
KW	GTP-binding; Lipoprotein.					
SQ	SEQUENCE 216 AA; 23955 MW; 3277D1EEDE77175B CRC64;					
Qy	19 SPPGSP--SYDLTGKVMVLGDVGKTCFLIQKDGAFLSGTTIATVGGIDFRNKVVTVDG 76					
Dd	::::					
Dd	3 APPARPADYDYLLKLLIGDSGVKSCLLRFSDGSFTT-SPTITIGIDFKIRTIEMDG 61					
Qy	77 VRVKLIQWDTAGQERFSRVTHAYRDQAQLLLLYDITNKSFDNIRAWLTEIHEYAQRDV 136					
Dd	:::					
Dd	62 KRILKLIQWDTAGQGRFTITAYRGAMGILLVDVTDESSFNINRNIRNIIEQHASNIV 121					
Qy	137 VIMLLGNKADM--SERVISEDGETLAREYGVPLETSARTKGNNVELAFLATAIKELKYRA 195					
Dd	:::					
Dd	122 NKILVGNKADMDSESKRAVPYSKQALADEYGIFFEASAKTNNVEVFSTAKDIKQL 181					
Qy	196 GH----QADEPSFIQIRDY-----VESOKKRSSCC 220					
Dd	: :					
Dd	182 AETDSKTEPOTIKINQEQAGTSAASQSGACC 214					
RESULT 8						
ID	Q40218	PRELIMINARY;	PRT;	214 AA.		
AC	Q40218;					
DT	01-NOV-1996 (TrEMBLrel. 01, Created)					
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)					
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)					
DE	RAB8D.					
OS	RAB8D.					
GN	Lotus japonicus.					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;					
OC	eurosid1 I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.					
ON	NCBI_TaxID=34305;					
OX	[1]					
RC	SEQUENCE FROM N.A.					
RP	TISSUE=ROOT NODELES.					
RC	MEDLINE=97231679; PubMed=9076991;					
RX	Borg S., Brandstrup B., Jensen T.J., Poulsen C.;					
RA	"Identification of new protein species among 33 different small GTP-					
RT	binding proteins encoded by cDNAs from Lotus japonicus, and expressed					
RT	of corresponding mRNAs in developing root nodules.";					
RL	Plant J. 11:237-250(1997).					
CC	-1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.					
DR	EMBL; Z73947; CAA98175.1; -;					
DR	HSP; P05713; 3RAB.					
DR	InterPro: IPR003579; GTPase_Rab.					
DR	InterPro: IPR001806; Ras.transfmg.					
DR	InterPro: IPR002078; Sig54.interact.					
DR	InterPro: IPR005225; Small_GTP.					
DR	Pfam: PF00071; ras; 1.					
DR	PRINTS; PR00449; RASTRNSFRMG.					
DR	SMART; SM00175; RAB; 1.					
DR	TIGRFAMS; TIGR00231; small_GTP; 1.					



```
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 214 AA; 23627 MW; 5E1A6B83505E50D7 CRC64;

Query Match 44.3%; Score 509.5; DB 10; Length 214;
Best Local Similarity 47.6%; Pred. No. 5.4e-39;
Matches 101; Conservative 45; Mismatches 55; Indels 11; Gaps 5;

QY 19 SPP--CSPSYDITGKVMVLGDTGVGKTCFLIQKDGAFLSGFTIATVGDIFRNKVVTVDG 76
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 3 APPARADYDYLKLLIGDGVGKSCLLRFSGSGFTT-SFTITIGIDFKIRTIELDG 61

QY 77 VYVKLIQWDTAGQERFSVTHAYYRDAQALLLYDITNKSFDNIRAWLTIETHEYAQRDV 136
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 62 KRVKLIQWDTAGQERFTITAYYRGAMGILLVYDVTDEAFNNTNRNLNRIEQHASDNV 121

QY 137 VIMLLGNKADM--SSERVIRSEDETALAREYGVVPFLETSAKTGMNVELAFALAIKELKYR- 194
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 122 NKILVGNKADMESKRAVPTSGQALADEYGIKFTSAKTNLNVVEVFFSIARDIKQRL 181

QY 195 --AGHQADEPSFQIRD----YVESQKKRSSCC 220
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 182 ADTDHKAEPITLTKINQDSAAAGERAANKSSCC 213

RESULT 9
Q949E2
ID Q949E2 PRELIMINARY; PRT; 203 AA.
AC Q949E2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Putative GTP-binding protein.
GN W455ERIPDK.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21329048; PubMed=11433398;
RA Mayer K., Murphy G., Tarchini R., Wambutt R., Volckaert G., Pohl T.,
RA Dueterhoeft A., Stiekema W., Entian K.D., Terryn N., Lemcke K.,
RA Haase D., Hall C.R., van Dodeweerdt A.M., Tingey S.V., Mewes H.W.,
RA Bevan M., Bancroft I.;
RT "Conservation of microstructure between a sequenced region of the
RT genome of rice and multiple segments of the genome of Arabidopsis
RT thaliana.";
RL Genome Res. 11:1167-1174(2001).
DR EMBL; AJ307662; CAC39050.1;
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding.
SQ SEQUENCE 203 AA; 22504 MW; B59DE515C09C5A82 CRC64;

Query Match 44.3%; Score 509; DB 10; Length 203;
Best Local Similarity 46.6%; Pred. No. 5.6e-39;
Matches 95; Conservative 49; Mismatches 52; Indels 8; Gaps 3;

QY 23 SPYSYDLTGKVMVLGDTGVGKTCFLIQKDGAFLSGFTIATVGDIFRNKVVTVDGVRVKLQ 82
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 2 NPEYDYLKLLIGDGVGKSCLLRFDSDYLD-SYISTIGVDFKIRTVQDGKTIKQLQ 60

QY 83 IWDTAGQERFSVTHAYYRDAQALLLYDITNKSFDNIRAWLTIETHEYAQRDVVIMLLG 142
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 61 IWDTAGQERFTITSSYRGAGIIVYDVTQDSFNQVNNKQWLNEDIRYASDNVKNLLVG 120

QY 143 NKADMSSERVIRSEDETALAREYGVVPFLETSAKTGMNVELAFALAIKELKYRAGHQ---- 198
```

```
Db 121 NKSDLTANKVSVSETAKAFADENGIPFMETSAKNATNVQEAFAAASIKDRMASQAAA 180
QY 199 -ADEPSFQIRDYVESQKKRSSCCS 221
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 181 NARPPTVQIRGQPVNOK--TSCCS 202

RESULT 10
Q40215
ID Q40215 PRELIMINARY; PRT; 216 AA.
AC Q40215;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE RAB8A.
GN RAB8A.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT NODULES;
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Braadstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
RT binding proteins encoded by cDNAs from Lotus japonicus, and expression
RT of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73944; CAA98172.1;
DR HSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 216 AA; 23776 MW; 569926CCA8D1B489 CRC64;

Query Match 44.3%; Score 509; DB 10; Length 216;
Best Local Similarity 45.2%; Pred. No. 6.1e-39;
Matches 103; Conservative 46; Mismatches 57; Indels 22; Gaps 5;

QY 1 MTGTPGAVATRDGEAPERSPPCSPSYDITGKVMVLGDTGVGKTCFLIQKDGAFLSGTFI 60
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 1 MAGAPAAARAD-----YDYLKLLIGDGVGKSCLLRFSGSGFTT-SFTI 46

QY 61 ATVGIDFRNKVVTVDGVRVKLIQWDTAGQERFSVTHAYYRDAQALLLYDITNKSFDN 120
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 47 TTIGIDFKIRTIELDGKRIKLIQWDTAGQERFTITAYYRGAMGILLVYDVTDEASFN 106

QY 121 IRAWLTIETHEYAQRDVVIMLLGNKADM--SSERVIRSEDETALAREYGVVPFLETSAKTMN 179
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 107 IKNNIRNTEQHASDNVNNKILVGNKADMESKRAVPTSGQALADEYGIKFTSAKTNLN 166

QY 180 VELAFALAIKELKYR---AGHQADEPSFQIRDYVES----QKKRSSCC 220
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 167 VEEVFFSIARDIKQRLADTDSRAEPQTIQINQPDASASGGQAQKSCC 214

RESULT 11
Q9SF91
ID Q9SF91 PRELIMINARY; PRT; 218 AA.
AC Q9SF91;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
```



Qy 25 SYDLTKVLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTVGVVRLQIW 84  
 Db 5 TYDLLFKLLIGDGVGKTCILFRSDDAFTS-TFISTIGIDFKIKTVELRGKKIKLQIW 63  
 Qy 85 DTAGQERFRSVTHAYYRDAQAALLLLYDITNKSSFNIRAWLTIEHYAQRDQVIMLGNK 144  
 Db 64 DTAGQERFRFTTSYRGAMGIMLYDITNKSSFNIRAWLTIEHYAQRDQVIMLGNK 123  
 Qy 145 ADMSSERVISGDTGLAREYGVPLFSAKTMVLAFLAIAKE-LKYRAGHQADPS 203  
 Db 124 CMTDKRVVNERGAIRAEHIREFWSAKSNINIERAFCELAAILDKTSGRESAENQ 183  
 Qy 204 FQIRDYVESQKRR---SSCCS 221  
 Db 184 ERVIDRRNQEKAPGYSKCCA 204  
 RESULT 13  
 Q40203  
 ID Q40203 PRELIMINARY; PRT; 202 AA.  
 AC Q40203;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE RAB1C.  
 GN RAB1C.  
 OS Lotus japonicus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.  
 OX NCBI\_TaxID=34305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-ROOT NODULES;  
 RX MEDLINE-97231679; PubMed-9076991;  
 RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;  
 RT "Identification of new protein species among 33 different small GTP-  
 binding proteins encoded by cDNAs from Lotus japonicus, and expression  
 of corresponding mRNAs in developing root nodules.";  
 RL Plant J. 11:237-250(1997).  
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; 273932; CAA98160.1; -;  
 DR HSSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase\_Rab.  
 DR InterPro; IPR001806; Ras\_trnsfrmng.  
 DR InterPro; IPR002078; Sig54\_interact.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMS; TIGR00231; small\_GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 202 AA; 22633 MW; 91633BBFAD07662F CRC64;  
 Query Match 44.1%; Score 507.5; DB 10; Length 202;  
 Best Local Similarity 46.8%; Pred. No. 7.6e-39;  
 Matches 95; Conservative 46; Mismatches 55; Indels 7; Gaps 3;  
 Qy 23 SPSYDLTKVLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTVGVVRLQ 82  
 Db 2 NPEYDLTKVLLIGDGVGKSCLLLRFDADSYLD-SYSTIGVDFKIRVEQDGKTIKIQ 60  
 Qy 83 IWDTAGQERFRSVTHAYYRDAQAALLLLYDITNKSSFNIRAWLTIEHYAQRDQVIML 142  
 Db 61 IWDTAGQERFRFTTSYRGAMGIMLYDITNKSSFNIRAWLTIEHYAQRDQVIML 120  
 Qy 143 NKADSSERVISGDTGLAREYGVPLFSAKTMVLAFLAIAKE-LKYRAGHQADPS 198  
 Db 121 NKCDITENKVSYSYETAKAFADIEIGIPFWETSNAKATNVEQAFMAAEIKRMASQPVNN 180  
 Qy 199 ADEPSFIQIRDYVESQKRRSCCS 221

Db 181 ARPPTVQIRGQPVNOK--SGCCS 201  
 RESULT 14  
 Q9LZD4  
 ID Q9LZD4 PRELIMINARY; PRT; 216 AA.  
 AC Q9LZD4;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE GTP-binding protein-like (Putative GTP-binding protein).  
 GN F12E4\_300.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,  
 RA Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,  
 RA Onodera C.S.; Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,  
 RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,  
 RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,  
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
 RA Southwick A.; Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Full length cDNA of gene F12E4\_300 (GI:7378637).";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,  
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,  
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,  
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,  
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,  
 RA Shinozaki K.; Davis R.W., Ecker J.R., Theologis A.;  
 RT "Full length cDNA of gene F12E4\_300 (GI:7378637).";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; AL162751; CAB83313.1; -;  
 DR EMBL; AY056331; AAK07200.1; -;  
 DR EMBL; AY035134; AAK059629.1; -;  
 DR HSSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase\_Rab.  
 DR InterPro; IPR001806; Ras\_trnsfrmng.  
 DR InterPro; IPR002078; Sig54\_interact.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMNG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMS; TIGR00231; small\_GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; UNKNOWN\_1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 216 AA; 24037 MW; 50C7D3AF4DC86F67 CRC64;  
 Query Match 44.1%; Score 507; DB 10; Length 216;  
 Best Local Similarity 47.5%; Pred. No. 9.4e-39;  
 Matches 103; Conservative 46; Mismatches 56; Indels 12; Gaps 4;  
 Qy 15 APERSPPCSPSYDLTKVLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTV 74  
 Db 15 APERSPPCSPSYDLTKVLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTV 74

```
Db 4 APARA---RSDYDLIKLLIGDSGVKSCLLLRFSDDTFTT-SFITTIGIDFKIRTVEL 59
QY 75 DGVVRKLIWDTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQR 134
Db 60 DGRKRIQLIWDTAGQERFRFTTAYYRGAMGILLVYDVTDESSEFNIRNMKNIEQHSD 119
QY 135 DVVIMLLGNKADM-SSEVRIRSEDEGTIAREYGVFPLETSAKTGMNVELAFIAIKELY 193
Db 120 NVNKLIVGNKADMEDESKRAVPTAKQALADEYGIKFETSAKTNLNVNFMVMSIAKDIQ 179
QY 194 R-----AGHQADEPSFOIRDYVESQKKRSSCCSEF 223
Db 180 RLUTETDKAEPQGIKITKQDTRAASSSTAESACCSIV 216

RESULT 15
Q8VWF9
ID Q8VWF9 PRELIMINARY; PRT; 216 AA.
AC Q8VWF9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ras-related protein RAB8-4 (Ras-related protein RAB8-2).
GN RAB8-4 OR RAB8-2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHT YELLOW 2;
RA Torimoto N., Shimada K., Ito K., Yamamoto K.;
RT "Characterization of Rab8 from tobacco BY-2 cell.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB079023; BAB84325.1; -.
DR EMBL; AB079021; BAB84323.1; -.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmng.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF0071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
SQ SEQUENCE 216 AA; 23971 MW; D5E87FABD0D49E13 CRC64;

Query Match 44.1%; Score 507; DB 10; Length 216;
Best Local Similarity 46.9%; Pred No. 9.4e-39;
Matches 100; Conservative 45; Mismatches 56; Indels 12; Gaps 4;

QY 19 SPP--CSPSYDLTGKYMLLGDTGVGKTCFLIOFKDGAFLSGTFIATVGIDFRNKVYVDG 76
Db 3 APPARADYDYLKLLIGDSGVKSCLLLRFSDDTFTT-SFITTIGIDFKIRTIELDG 61
QY 77 VRVKLIWDTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQRDV 136
Db 62 KRKLIQIWDTAGQERFTTAYYRGAMGILLVYDVTDESSEFNIRNMKNIEQHSDNV 121
QY 137 VIMLLGNKADM-SSEVRIRSEDEGTIAREYGVFPLETSAKTGMNVELAFIAIKELKYA 195
Db 122 NKILVGNKADMEDESKRAVPTSKQALADEYGIKFETSAKTNNNVNVEVFFSIARDIKRL 181
QY 196 GHQADEPSFOIRDYVESQK-----KRSSCC 220
Db 182 SEDSKTEPAIRINQSDQAGTSGQAQKSSCC 214
```

Search completed: November 17, 2002, 16:41:51  
Job time : 82 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 16:36:38 ; Search time 43 Seconds  
(without alignments)  
498.558 Million cell updates/sec

Title: US-09-817-199A-2

Perfect score: 1150

Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDYVESOKRRSSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	742	64.5	190	2 JC2528	GTP-binding protei
2	533	46.3	196	2 T15123	hypothetical prote
3	517.5	45.0	203	2 B34716	GTP-binding protei
4	516	44.9	206	2 I78851	GTP-binding protei
5	513.5	44.7	207	2 B49647	GTP-binding protei
6	513.5	44.7	207	2 B36364	GTP-binding protei
7	512.5	44.6	215	2 T14565	GTP-binding protei
8	507	44.1	216	2 T48378	GTP-binding protei
9	507	44.1	258	2 B86153	ARA-5 [imported] -
10	506	44.0	216	2 S33900	GTP-binding protei
11	505.5	44.0	201	2 T28971	hypothetical prote
12	500	43.5	202	2 S38740	GTP-binding protei
13	500	43.5	216	2 T45901	GTPase ATRAB8 - Ar
14	498	43.3	208	2 A34716	GTP-binding protei
15	497.5	43.3	202	2 S41450	GTP-binding protei
16	497.5	43.3	209	2 B38625	GTP-binding protei
17	497.5	43.3	215	2 S57478	GTP-binding protei
18	495	43.0	203	2 B38202	GTP-binding protei
19	495	43.0	216	2 S57471	GTP-binding protei
20	495	43.0	216	2 J50640	GTP-binding protei
21	494	43.0	200	2 D36384	GTP-binding protei
22	490.5	42.7	200	2 A38625	GTP-binding protei
23	490.5	42.7	215	2 S57462	GTP-binding protei
24	489	42.5	222	2 T14405	small GTP-binding
25	488	42.4	203	2 S34253	GTP-binding protei
26	487	42.3	200	2 S12790	GTP-binding protei
27	486	42.3	200	2 B42148	GTP-binding protei
28	486	42.3	202	2 S72515	GTP-binding protei
29	486	42.3	203	2 JC4105	GTP-binding protei

30	485	42.2	196	2 PS0279	GTP-binding protei
31	484.5	42.1	215	2 S57474	GTP-binding protei
32	483	42.0	221	2 H71444	GTP-binding protei
33	481.5	41.9	224	2 T33855	hypothetical prote
34	479.5	41.7	204	2 JC7589	Sec4p homolog - ye
35	479.5	41.7	208	2 A38202	GTP-binding protei
36	478.5	41.6	206	2 T14391	GTP-binding protei
37	478	41.6	201	2 S06147	GTP-binding protei
38	478	41.6	203	2 JC1247	GTP-binding protei
39	477.5	41.5	205	2 S38339	GTP-binding protei
40	476.5	41.4	205	2 T33781	hypothetical prote
41	476	41.4	217	2 S36365	GTP-binding protei
42	473.5	41.2	203	2 A49647	GTP-binding protei
43	468.5	40.7	201	2 S39565	GTP-binding protei
44	463.5	40.3	203	2 S51495	GTP-binding protei
45	459.5	40.0	203	2 S30096	GTP-binding protei

ALIGNMENTS

RESULT 1

JC2528

GTP-binding protein Rab26 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 02-Feb-2001

C:Accession: JC2528

R:Wagner, A.C.C.; Strowski, M.Z.; Goetze, B.; Williams, J.A.

Biochem. Biophys. Res. Commun. 207, 950-956, 1995

A:Title: Molecular cloning of a new member of the Rab protein family, Rab26, from rat

A:Reference number: JC2528; MUID:95169156; PMID:7864900

A:Accession: JC2528

A:Molecule type: mRNA

A:Residues: 1-190 <WAG>

A:Cross-references: GB:T08521

A:Note: The authors translated the codon GTG for residue 49 as Leu, GAA for residue 1

C:Genetics:

A:Gene: Rab26

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; nucleotide binding; P-loop

F:4-11/Region: nucleotide-binding motif A (P-loop)

F:52-58/Region: GTP binding #status predicted

F:111-114/Region: GTP binding #status predicted

F:139-143/Region: GTP binding #status predicted

Query Match 64.5%; Score 742; DB 2; Length 190;  
Best Local Similarity 72.9%; Pred. No. 3.4e-57;  
Matches 137; Conservative 28; Mismatches 23; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLQFDGAFSLGTFATVGDIFRNKVVTVDGVVRKLIQIWDTAGQERF 92

Db 1 MLVGSVGKTKLLVRFKDGAFAGTFTSTVGDIFRNKVLVDVGKVKVQIWDTAGQERF 60

QY 93 RSVTHAYRDAQALLLLYDITNKSSFDNTRAWLTFEHEYAQRDVVIMLGNKADMSERV 152

Db 61 RSVTHAYRDAHALLLYDITNKSSFDNTRAWLTFEHEYAQRDVVIMLGNKVDTSQERV 120

QY 153 IRSEGEFIALREYGVPEFLETSKATGMNVELAFATLAKELKYRAGHQADEPSFQIRDYVES 212

Db 121 VKREGEKLAKEYGLPFMETSAKSLGNLVDAFTATAKELKQRSTKAPSEPRFLHLDYVKR 180

QY 213 QKKRSSCC 220

Db 181 EGRGVSCC 188

RESULT 2

T15123

hypothetical protein W01H2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000

C:Accession: T15123

R:Minx, P.; Wohldmann, P.

submitted to the EMBL Data Library, April 1997  
A:Description: The sequence of C. elegans cosmid W01H2.  
A:Reference number: Z18296

A:Accession: T15123  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-196 <MIN>

A:Cross-references: EMBL:AF000192; NID:g1946982; PID:g1946985; PIDN:AAB52888.1; GSPDB:GN  
A:Experimental source: strain Bristol N2; clone W01H2

C:Genetics:  
A:Gene: CESP:W01H2.3

A:Map position: X  
A:Introns: 41/3; 55/3; 95/3; 134/3; 160/2

C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
F:3-119/Domain: translation elongation factor Tu homology <ETU>

Query Match 46.3%; Score 533; DB 2; Length 196;

Best Local Similarity 54.5%; Pred. No. 5.2e-39;

Matches 109; Conservative 34; Mismatches 41; Indels 16; Gaps 5;

QY 31 KYVLLGDTGVGKTCFLIQKDAFLSGTFTATVGVDFRNKVVTVGVRVKLQIWDTAGOE 90

Db 4 KYVLLGDSCTGKTCCLIRKDKAFLNNFISVGVDFRNKLTITGDKKVKLQIWDTAGOE 63

QY 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIHEYAQRDVVIMLLGNKADMSSE 150

Db 64 RFRSVTTSYVRADALLLYDITANRASFCNWLISQIKEYGKAQVTVLVGNKCDL--P 121

QY 151 RVIRSDGETLAREYGVPLETSAGTKGMVVELAFIAIAK---ELKY---RAGHQADEPSF 204

Db 122 RAVPTDEGRRLAEAYQIPMETSAKTGFNVDFRAFLGLAERMLKLYGFVPGGEMADTIS- 180

QY 205 QIRDYVESQKR--SSCCSF 222

Db 181 ----VADTKKPEIARCTCF 195

#### RESULT 3

B34716

C:Superfamily: Dictyostellium discoideum

C:Keywords: Dictyostellium discoideum

C:Date: 29-Jun-1990 #sequence\_revision 29-Jun-1990 #text\_change 02-Feb-2001

C:Accession: B34716; B61571

R:Saxe, S.A.; Kimmel, A.R.  
Mol. Cell. Biol. 10, 2367-2378, 1990

A:Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostellium discoideum with sequ

A:Reference number: A34716; MUID:90220623; PMID:2109188

A:Accession: B34716

A:Molecule type: DNA

A:Residues: 1-203 <SAX>

A:Cross-references: GB:M34457

R:Saxe, S.A.; Kimmel, A.R.  
Dev. Genet. 9, 259-265, 1988

A:Title: Genes encoding novel GTP-binding proteins in Dictyostellium.

A:Reference number: A61571; MUID:89209367; PMID:3149563

A:Accession: B61571

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-203 <SA2>

C:Genetics:  
A:Gene: SAS2

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F:16-131/Domain: translation elongation factor Tu homology <ETU>  
F:22-29/Region: nucleotide-binding motif A (P-loop)  
F:128-131/Region: GTP-binding NKXD motif  
F:158-160/Region: GTP-binding SAK/L motif  
F:202,203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match

Best Local Similarity 45.0%; Score 517.5; DB 2; Length 203;

Matches 97; Conservative 50; Mismatches 54; Indels 5; Gaps 3;

QY 15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQKDAFLSGTFTATVGVDFRNKVVTV 74

Db 3 SPATNKPA--AYDFLVKLLIGDSGVGSKLLRSDGSF--TPSFATIGIDFKIRTIEL 59

QY 75 DGVRVKLOIWDTAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIHEYAQR 134

Db 60 EGKRILQIWDTAGOERFTITAYRGAMGILLVDYTDKSFSGSIRNWNINIQHASD 119

QY 135 DVVIMLLGNKADMSRVSIRSDGETLAREYGVPLETSAGTKGMVVELAFIAIAKELKYR 194

Db 120 SVNKMILGNKCDMTEKKVVDSSRGKSLADEYGIKLETSAKNSVNVEEAFICLAKDKKR 179

QY 195 AQHQADEPSFQIRDYVESQKRSSCC 220

Db 180 MIDTPNDPDHTI--CITPNKKNTCC 203

#### RESULT 4

I78851

C:Superfamily: GTP-binding protein MEL - mouse

N:Alternate names: gene MEL protein

C:Species: Mus sp. (mouse)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 02-Feb-2001

C:Accession: I78851

R:Nimmo, E.R.; Sanders, P.G.; Padua, R.A.; Hughes, D.; Williamson, R.; Johnson, K.J.

Oncogene 6, 1347-1351, 1991

A:Title: The MEL gene: a new member of the RAB/YPT class of RAS-related genes.

A:Reference number: I58355; MUID:91360267; PMID:1886711

A:Accession: I78851

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-206 <RES>

A:Cross-references: GB:S53270; NID:g234747; PIDN:AAB19682.1; PID:g234748

C:Genetics:

A:Gene: MEL

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.9%; Score 516; DB 2; Length 206;

Best Local Similarity 49.5%; Pred. No. 1.7e-37;

Matches 101; Conservative 41; Mismatches 56; Indels 6; Gaps 3;

QY 25 SYDLTGKVMLLGDTGVGKTCFLIQKDAFLSGTFTATVGVDFRNKVVTVGVRVKLQIW 84

Db 4 TYDYLFKLLIGDSGVGKTCVLFSEDAF--NSTFSTIGIDFKIRTIEDGKRILQIW 62

QY 85 DTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIHEYAQRDVVIMLLGNK 144

Db 63 DTAGQERFTITAYYRGAMGIMLVYDITNEKSFSDNIRNWNIRNIEHASADVEKMILGNK 122

QY 145 ADMSSRVSIRSDGETLAREYGVPLETSAGTKGMVVELAFIAIAKELKYRAGHQ----AD 200

Db 123 CDVNDKRVQSKERGEKALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKNWKATAA 182

QY 201 EPSFQIRDYVESQKRSSC--CSFM 223

Db 183 GSHGVKRVITVEQKRISFFRCSLL 206

#### RESULT 5

B49647

C:Superfamily: GTP-binding protein rab8 - human

C:Species: Homo sapiens (man)

C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Jan-2001

C:Accession: B49647; S36817

R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavitian, A.; Lou

J. Cell Biol. 124, 101-115, 1994

A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c

A:Reference number: A49647; MUID:94124602; PMID:8294494

A:Accession: B49647

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-207 <ZAR>

A:Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA40065.1; PID:g452318

R:Joberty, G.; Tavittian, A.; Zahraoui, A.

FEBS Lett. 330, 323-328, 1993

A:Title: Isoprenylation of Rab proteins possessing a C-terminal CaaX motif.

A:Reference number: S36817; MUID:93387463; PMID:8375503

A:Accession: S36817

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 175-186 <JOB>

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFIATVGDIFRNKVVTVGVRVKLQIW 84

Db 4 TYDYLKLLIGDSGVGKTCVLFRESEDAF-NSTPISITIGIDFKIRTIELDGKRIKLQIW 62

Qy 85 DTAGQERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNK 144

Db 63 DTAGQERPTTITAYRGAMGIMLVYDITNEKSFNIRNIRNIEHASADVEKMILGNK 122

Qy 145 ADMSSERVISEDGETLAREYGVPPLETSAKTGMNVELAFIAIAKELKYRAGHQAD---- 200

Db 123 CDVNDKRVKSRGKGLDYGKIFMETSAKININVENAFETLARDIKAKMDKKLEGNSP 182

Qy 201 EPSFOIRDYVESOKRSS 218

Db 183 QGSNOGVKITPDQQRSS 200

RESULT 6

B36364

GTP-binding protein rab8 - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 02-Feb-2001

C:Accession: B36364; S15604

R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.

Mol. Cell. Biol. 10, 6578-6585, 1990

A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.

A:Reference number: A36364; MUID:91061765; PMID:2123294

A:Accession: B36364

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-207 <CHA>

A:Cross-references: GB:X56385; NID:g920; PIDN:CAB56776.1; PID:g6006436

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFIATVGDIFRNKVVTVGVRVKLQIW 84

Db 4 TYDYLKLLIGDSGVGKTCVLFRESEDAF-NSTPISITIGIDFKIRTIELDGKRIKLQIW 62

Qy 85 DTAGQERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNK 144

Db 63 DTAGQERPTTITAYRGAMGIMLVYDITNEKSFNIRNIRNIEHASADVEKMILGNK 122

Qy 145 ADMSSERVISEDGETLAREYGVPPLETSAKTGMNVELAFIAIAKELKYRAGHQAD---- 200

Db 123 CDVNDKRVKSRGKGLDYGKIFMETSAKININVENAFETLARDIKAKMDKKLEGNSP 182

Qy 201 EPSFOIRDYVESOKRSS 218

Db 183 QGSNOGVKITPDQQRSS 200

RESULT 7

TL4565

GTP-binding protein - beet

N:Alternate names: small G protein

C:Species: Beta vulgaris (beet)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000

C:Accession: TL4565

R:Dallery, E.; Quief, S.; Ben Jilany, K.E.; Kerckaert, J.; Hagege, D.

submitted to the EMBL Data Library, May 1995

A:Description: Molecular cloning and structural analysis of cDNAs that encode three s

A:Reference number: Z18142

A:Accession: TL4565

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-215 <DAL>

A:Cross-references: EMBL:Z49152; NID:g974775; PID:g974776

A:Experimental source: strain D100 KS 38080

C:Function:

A:Description: GTP-binding

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFIATVGDIFRNKVVTVDG 76

Db 3 APPARADYDYLIKLLIGDSGVGKSCLLLRFSGSEFT-SFITTIGIDFKIRTIELDG 61

Qy 77 VRVQLQIWDTAGQERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQRDV 136

Db 62 KRILQIWDTAGQERPTTITAYRGAMGIMLVYDITDESSFNIRNIRNIEHASADNV 121

Qy 137 VIMLLGNKADM-SSERVISEDGETLAREYGVPPLETSAKTGMNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMESKRAVPTAKQALADYEGIKFFETSAKTNLNVVEVFFSIARDIKQRL 181

Qy 196 G----HQADEPSFQIRDYVESOKK---RSSCC 220

Db 182 ADSDTROEAQPSITIKPADQSGNQAAKSACC 213

RESULT 8

T48378

GTP-binding protein-like - Arabidopsis thaliana

N:Alternate names: protein F12E4.300

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000

C:Accession: T48378

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke,

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24492

A:Accession: T48378

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-216 <BEV>

A:Cross-references: EMBL:AL162751

A:Experimental source: cultivar Columbia; BAC clone F12E4





[illegible]

RESULT 12  
S38740  
GTP-binding protein - rice  
N:Alternate names: ras-related small GTP-binding protein  
C:Species: *Oryza sativa* (rice)  
C:date: 27-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 02-Feb-2001  
C:Accession: S38740  
R:Kidou, S.; Anal, T.; Umeda, M.; Aotsuka, S.; Tsuge, T.; Kato, A.; Uchimiya, H.  
FEBS Lett. 332, 282-286, 1993  
A:title: Molecular structure of ras-related small GTP-binding protein genes of rice plant  
A:Reference number: S38740; MUID:94009718; PMID:8405471  
A:Accession: S38740  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-202 <KID>  
A:Cross-references: GB:S66160; NID:9432606; PIDN:AB28535.1; PID:9432607  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:9-124/Domain: translation elongation factor Tu homology <ETU>  
F:15-22/Region: nucleotide-binding motif A (P-loop)  
F:121-124/Region: GTP-binding NKXD motif  
F:151-153/Region: GTP-binding SAK/L motif  
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys Asp, Ser) #sta

Query Match	43.5%;	Score 500;	DB 2;	Length 202;
Best Local Similarity	46.3%;	Pred. No. 4e-36;		
Matches	94;	Conservative 49;	Mismatches 54;	Indels 6; Gaps
QY	23	SPSYDLTGKVMLLDGTGVGKTCFLIOFKDGAFLSGTFIATVGIDFRKNKVVTVDVGVRVKLQ	82	
Db	2	NPEYDYLKLLLLDGSVGKSCILLRFADDSYLE-SYISTIGVDYFKIRVEQDGKTIKQLQ	60	
QY	83	IWDTAGQERPRSYTHAYRDAQALLLLYDTNKSFDNTIRAWTETHEHYAQRDVIMLLG	142	
Db	61	IWDTAGQERPTTSSYIRGAHGIIIVYVDVTQGESFNPNKQWLEINDRYASENNVKNLLVG	120	
QY	143	NKADMSSYIRSEDEGTFLAREYGVFPLETSAKTGMNVELAFIAIAKEUKYIRAGHO----	198	
Db	121	NKCDLAENRVVSYEAGKALADEIGIFPLETSAKDATNVEKAFMTMAGEIKNRMASOGRTN	180	
QY	199	ADEPFSFOIRDYVESQKRRSCCS	221	
Db	181	ASKPA-TVOMPRFVAOOSCCS	202	

RESULT 13  
T45901  
GTPase ATRAB8 - Arabidopsis thaliana  
N:Alternate names: protein F4P12.310  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 02-Sep-2000  
C:Accession: T45901  
R:Bloeker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23016  
A:Accession: T45901  
A:Status: preliminary

A: Molecule type: DNA  
A: Residues: 1-216 <BLO>  
A: Cross-references: EMBL:AL132966  
A: Experimental source: cultivar Columbia; BAC clone F4P12  
C: Genetics:  
A: Map Position: 3  
A: Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3  
A: Note: F4P12.310  
C: Superfamily: ras transforming protein; translation elongation factor Tu homology  
F; 16-131/domain: translation elongation factor Tu homology <EU>

```

Query Match      43.5%; Score 500; DB 2; Length 216;
Best Local Similarity 46.5%; Pred. No. 4.3e-35;
Matches 99; Conservative 45; Mismatches 57; Indels 12; Gaps 5;

Qy 19 SPP--CSPSYDLTKCVMLLGDVGKTCFLQFKDGAFLSCFTIATGIDERNKVVVDG 76
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 3 APPARADYDYLKLLLLIGDVGKSCLLLRFDGSEFTT-SFTITGIDFKPIETLDG 61

Qy 77 VRVKLIQINDTAGQRRFRSVTHAYYRDQAQLLLLDITNKSFDNIRAWLTIHEYAQRDV 136
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 62 KRILQIWDTAGQRRFRTITTAYYRGAMGILLVDVDTDESFNNIRNWRNIEQHASDV 121

Qy 137 VIMLLGNKADM--SSEVRIRSEDTGELAREYGVPFLETSAKTGMVVELAFLAIKELKYRA 195
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 122 NKLIIVGNKADMDESRRAVPKSGQALADEYGNKFFETSAKTNLNVVEVFFSIADIKORL 181

Qy 196 GH-----QADEPSPFIQIRDY-----VESQKRKRSCC 220
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 182 ADTDARAPPQPIKINOSDQAGTSGATSOATOKSACC 214

```

RESULT 14  
A34716  
GTP-binding protein SAS1 - slime mold (Dictyostellium discoideum)  
C:Species: Dictyostellium discoideum  
C:Date: 29-Jun-1990 #sequence\_revision 29-Jun-1990 #text\_change 02-F-  
C:Accession: A34716; A61571  
R:Saxe, S.A.; Kimmel, A.R.  
Mol. Cell. Biol. 10, 2367-2378, 1990  
A:Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostellium d  
A:Reference number: A34716; MUID:50220623; PMID:2109188  
A:Accession: A34716  
A:Molecule type: mRNA  
A:Residues: 1-208 <SAX>  
A:Cross-references: GB:M34456  
R:Saxe, S.A.; Kimmel, A.R.  
Dev. Genet. 9, 259-265, 1988  
A:Title: Genes encoding novel GTP-binding proteins in Dictyostellium.  
A:Reference number: A61571; MUID:89209367; PMID:3149563  
A:Accession: A61571  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-208 <SA2>  
C:Genetics:  
A:Gene: SAS1

	Query Match	43.3%	Score 498	DB 2	Length 208
	Best Local Similarity	46.6%	Pred. No. 6.2e-36		
	Matches 95	Conservative 49	Mismatches 50	Indels 10	Gaps 3
Qy	23	SPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFSGCTFIATGVGIDFRNKVWTVGDVVRVQLQ	82		
	1				
Db	9	SAAYDYLKLLIGDSGVGKCLLLRFSEDSF-TPSFTTTIGIDPKETIELEGRRIKLQ	67		
Qy	83	IWDYTAGQRRFSRVTHAYYVRDQAQLLLLYDINFKSSFDNIRAWLTETIHEYAQRDVVIMLLG	142		

```
|||||
Db 68 IWDTAGQERFTTTAYIRGAMGILLVDYDEKSFNGINRNWIRNIEQHAATDSVKNKMLIG 127
QY 143 NKADMSERVIRSEDGETLAREYGVFPFLETSAKTGMNVELAFLAIKELKYRAGHQADEP 202
Db 128 NKCDMAEKVVDSRSGSLADEYIGKFLTSAKNSINVEAFISLAKDIKKR---MIDTP 184
QY 203 SFQIR-----DYVESQKRSSCC 220
Db 185 NEQPQVVQPGTNLGANNKKKACC 208

RESULT 15
S41430
GTP-binding protein, ras-like (clone vfa-ypt1) - fava bean
N:Alternate names: guanine nucleotide regulatory protein
C:Species: Vicia faba (fava bean)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 02-Feb-2001
C:Accession: S41430
R:Saalbach, G.; Thielmann, J.
submitted to the EMBL Data Library, January 1994
A:Description: Sequences of cDNA clones from cotyledons of Vicia faba encoding ypt/rab-x
A:Reference number: S41430
A:Accession: S41430
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-202 <SAA>
A:Cross-references: EMBL:295590; NID:g452358; PIDN:CAA82707.1; PID:g452359
A:Experimental source: clone vfa-ypt1
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
F:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKXD motif
F:151-153/Region: GTP-binding SAK/L motif
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
```

```
Query Match 43.3%; Score 497.5; DB 2; Length 202;
Best Local Similarity 45.8%; Pred. No. 6.6e-36;
Matches 93; Conservative 46; Mismatches 57; Indels 7; Gaps 3;

QY 23 SPYSYDLTGKVMILGDTGVGKTCFLIQKGAFLSGTFIATVGDIFRNKVVTVGVRVKLQ 82
Db 2 NPEYDYLKLLIGDSGVGKSCLLRFADDSYID-SYISTIGVDFKIRTVQDQGTIKLQ 60
QY 83 IWDTAGQERFRSVTHAYRDAQALLLYDITNKSSFNIRAWLTEIHEYAQRDVIWMLIG 142
Db 61 IWDTAGQERFTTSSYIRGAHGLIIVDVTDEESFNNKQWLSEIDRYASDNVKNLLVG 120
QY 143 NKADMSERVIRSEDGETLAREYGVFPFLETSAKTGMNVELAFLAIKELKYRAGHQ--- 198
Db 121 NKCDLTENRAVPYETAKAFADAEIGIPMETSAKDSTNVEQAFMAMASSIKERMASOPTNN 180
QY 199 ADEPSPQIRDYVESQKRSSCCS 221
Db 181 ARPPTVQIRGQPVGQK--SGCCS 201
```

Search completed: November 17, 2002, 16:42:55  
Job time : 44 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 17, 2002, 16:42:58 ; Search time 52 seconds  
(without alignments)  
1584.639 Million cell updates/sec

Title: US-09-817-199A-2  
Perfect score: 1150  
Sequence: 1 MTGPGAVATDGEAPERSP.....FQIRDYVESQKRSSCSFM 223

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 328367 seqs, 184756068 residues

Total number of hits satisfying chosen parameters: 656734

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO\_spool/US09817199/runat\_13112002\_134903\_5200/app\_query.fasta\_1.391  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09817199.ecgn.1.1.20.erunat.13112002.134903.5200  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:\*

1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2.6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2.6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2.6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	100.0	2674	10	US-09-817-199A-1
2	1145	99.6	1116	10	US-09-794-257-13
3	1140	99.1	875	12	US-10-051-986-10
4	977	85.0	576	10	US-09-794-257-15

5	720	62.6	447	10	US-09-867-550-1811	Sequence 1811, Ap
6	511.5	44.5	2497	10	US-09-834-975-879	Sequence 879, App
7	511.5	44.5	2497	10	US-09-834-975-885	Sequence 885, App
8	511.5	44.5	2497	10	US-09-834-975-894	Sequence 894, App
9	511.5	44.5	2497	10	US-09-834-975-896	Sequence 896, App
10	507	44.1	881	10	US-09-770-445-529	Sequence 529, App
11	506	44.0	1537	10	US-09-925-300-631	Sequence 631, App
12	501	43.6	624	10	US-09-794-257-9	Sequence 9, Appli
13	501	43.6	1161	10	US-09-794-257-7	Sequence 7, Appli
14	500	43.5	651	9	US-09-938-842A-836	Sequence 836, App
15	496.5	43.2	609	9	US-09-938-842A-832	Sequence 832, App
16	483	42.0	1274	10	US-09-925-302-91	Sequence 91, Appli
17	470	40.9	925	10	US-09-967-736-4	Sequence 4, Appli
18	448.5	39.0	639	10	US-09-350-874-66	Sequence 66, Appli
19	435	37.8	3257	10	US-09-817-198A-1	Sequence 1, Appli
20	420.5	36.6	896	10	US-09-770-445-478	Sequence 478, App
21	413.5	36.0	3124	10	US-09-925-302-340	Sequence 340, App
22	411	35.7	894	10	US-09-770-445-487	Sequence 487, App
23	408	35.5	1061	10	US-09-880-107-3393	Sequence 3393, Ap
24	401.5	34.9	585	10	US-09-917-800A-1461	Sequence 1461, Ap
25	399.5	34.7	771	10	US-09-728-445-652	Sequence 652, App
26	392.5	34.1	654	9	US-09-938-842A-2113	Sequence 2113, Ap
27	386	33.6	2771	10	US-09-834-765-1	Sequence 1, Appli
28	385	33.5	857	10	US-09-917-800A-1426	Sequence 1426, Ap
29	383	33.3	538	10	US-09-924-035A-794	Sequence 794, App
30	383	33.3	1042	10	US-09-954-456-563	Sequence 563, App
31	380	33.0	3936	10	US-09-919-172-49	Sequence 49, Appli
32	376	32.7	844	10	US-09-770-445-646	Sequence 646, App
33	374.5	32.6	412	10	US-09-960-352-1218	Sequence 1218, A
34	374.5	32.6	676	10	US-09-770-149-374	Sequence 374, App
35	373.5	32.5	642	10	US-09-794-257-6	Sequence 6, Appli
36	373.5	32.5	1023	10	US-09-794-257-4	Sequence 4, Appli
37	372	32.3	803	12	US-10-051-986-13	Sequence 13, Appli
38	371	32.3	4083	10	US-09-817-182-1	Sequence 1, Appli
39	370.5	32.2	645	9	US-09-938-842A-774	Sequence 774, App
40	370.5	32.2	900	10	US-09-770-445-469	Sequence 469, App
41	369	32.1	645	9	US-09-938-842A-74	Sequence 74, Appli
42	368	32.0	1007	10	US-09-822-849A-563	Sequence 563, App
43	368	32.0	1022	10	US-09-920-300A-1703	Sequence 1703, Ap
44	368	32.0	1022	12	US-10-033-528-1703	Sequence 1703, Ap
45	367.5	32.0	2251	10	US-09-834-765-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-817-199A-1

; Sequence 1, Application US/09817199A

; Patent No. US20020142380A1

; GENERAL INFORMATION:

; APPLICANT: SHAO, Wei et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01187

; CURRENT APPLICATION NUMBER: US/09/817,199A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2674

; TYPE: DNA

; ORGANISM: Human

US-09-817-199A-1

Alignment Scores:

Pred. No.:	4,11e-136	Length:	2674
Score:	1150.00	Matches:	223
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-817-199a-2 (1-223) x US-09-817-199a-1 (1-2674)

```
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 42 ATGACGGGACGCCAGCGCGCTTGCCACCGGGATGGGAGGCCCGCCGAGCGCTCCCG 101
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 102 CCTGTCAGTCCGAGCTACGACCTACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 161
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 162 GGCAAAACATGTTCTCTGATCAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCATCA 221
QY 61 AlaThrValGlyLysPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 281
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 282 CTGCAGATCTGGGACACCGCTGGGCAGGAACGGTTCGGAAGCGTCACCCATGCTTATTAC 341
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 342 AGAGATGCTCAGGCGCTTGCTCTGCTGTATGACATCACCACAAATCTTCTTCGACAAAC 401
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 402 ATCAGGCGCTGGCTCACTGAGATTCATGATGATGCCAGGAGCGTGTGTATGATGCTG 461
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 462 CTAGGCAACAAGCGGATATGAGCAGCAAGAGTATCGCTTCCGAAGACGGAGAGACC 521
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db 522 TTGGCAGGAGGTACGGGTGCTCCCTCTGGAGACCGAGCGCCCAAGACTGGCATGAATGTG 581
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db 582 GAGTTAGCCTTTCTGGCCATCGCCAAAGAACTGAATATCCGGCGCGGCATCAGGCGGAT 641
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 642 GAGCCCGAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAGCGCTCCAGCTGCTGC 701
QY 221 SerPheMet 223
Db 702 TCCTTCATG 710
```

## RESULT 2

```
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794, 257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)...(699)
US-09-794-257-13
```

## Alignment Scores:

```
Pred. No.: 4,92e-136 Length: 1116
Score: 1145.00 Matches: 222
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.57% Indels: 0
DB: 10 Gaps: 0
```

US-09-817-199a-2 (1-223) x US-09-794-257-13 (1-1116)

```
QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
Db 31 ACGGGACGCCAGCGCGCTTGCCACCGGGATGGGAGGCCCGCCGAGCGCTCCCGGCC 90
QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 91 TCCAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGC 150
QY 42 LysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db 151 AAAACATGTTCTCTGATCAATTCAAAGACGGGGCTTCTCTCGGAACCTTCATAGCC 210
QY 62 ThrValGlyLysPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 211 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGATGAAGCTG 270
QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 271 CAGATCTGGACACCGCTGGGAGGACCGTTCGAGAGCTCACCATGCTTATACAGA 330
QY 102 AspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 331 GATGCTCAGCGCTTGCTCTGCTGTATGACATCACCACAAATCTTCTTTCGACAAACATC 390
QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
Db 391 AGGGCTTGCTCACTGAGATTCATGATGCCAGGAGCGTGGTGATCATGCTGCTGCTA 450
QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db 451 GGCAACAAGCGGATATGAGCAGCGAAGAGTATCCGTTCCGAGACCGAGAGACCTTG 510
QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db 511 GCGAGGAGTACGGTGTCTCCCTTCTGGAGACCGACCGCCCAAGACTGGCATGAATGTGGAG 570
QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db 571 TTAGCCTTTCTGGCCATCCCAAGGAACCTGAATATCCGGCGCGGCATCAGCGGGATGAG 630
QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
Db 631 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAAGCGCTCCAGCTGCTGCTCC 690
QY 222 PheMet 223
Db 691 TTCATG 696
```

## RESULT 3

```
US-10-051-986-10
; Sequence 10, Application US/10051986
; Patent No. US20020146770A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Tang, Y. Tom
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; Batra, Sajeev
; Baugho, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
```



Db 241 ACCAACAATCTCTTTCCACACATCAGGGCTGCTCCTCAGATTATGAGTAGTCC 300  
QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152  
Db 301 CAGAGGAGCGTGGTATCATGCTAGTACCAACAGCGGATATGAGCAGCGAAGAGTG 360  
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172  
Db 361 ATCCGTTCCGAAGACGAGAGACCTTGGCCAGGAGTAGCGTGTCTCCCTTCCTGGAGACC 420  
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192  
Db 421 AGCGCCAAAGACTGGCATGAATGTGGAGTTAGCCTTCTGGCCATGCCCAAGGAACCTGAAA 480  
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212  
Db 481 TACCGGGCCGGGCATCCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGAGAGTCC 540  
QY 213 GlnLysLysArgSerSerCysSerPheMet 223  
Db 541 CAGAAGAAGCGCTCCAGCTGCTCCTTCATG 573

## RESULT 5

US-09-867-550-1811  
; Sequence 1811, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1811  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)  
; OTHER INFORMATION: Wherein n is one of a or t or c or g  
US-09-867-550-1811

Alignment Scores:  
Pred. No.: 11e-82 Length: 447  
Score: 720.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.61% Indels: 0  
DB: 10 Gaps: 0

US-09-817-199a-2 (1-223) x US-09-867-550-1811 (1-447)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
Db 34 ATCAGCGGACGCCAGCGCGCTTGCACCCGGATGGCGAGSCCCCGAGCGCTCCCG 93  
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
Db 94 CCCTGCAGTCCGAGCTACGACCTTCACGGGCAAGGTATGCTTCTGGGAGACACAGCGCTC 153  
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
Db 154 GCGAACAACATGTTCTCTGATGCCAATTCAAAGACGGGCGCTTCTCTGTCGGGAACCTTCATA 213

QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
Db 214 GCCACCGTCGGATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGAGTGAAG 273  
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
Db 274 CTGCAAGATCTGGACACCCCTGGCAGGAAACGGTCCGGAAGCGTCACCCATGCTTATTAC 333  
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
Db 334 AGAGATGCTCAGSCCTTGCTTCTGCTGATGATCATCACCAACAATCTCTTCGACAAC 393  
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValle 138  
Db 394 ATCAGGCGCTGCTCCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 447

## RESULT 6

US-09-834-975-879  
; Sequence 879, Application US/09834975  
; Patent No. US20020110815A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Brown, Jeffrey  
; APPLICANT: Bolt, Andrew  
; APPLICANT: Van Huffel, Christophe  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS  
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
; TITLE OF INVENTION: OF HUMAN CANCERS  
; FILE REFERENCE: MRI-016B  
; CURRENT APPLICATION NUMBER: US/09/834,975  
; CURRENT FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/197,538  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 1046  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 879  
; LENGTH: 2497  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(2497)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-834-975-879

Alignment Scores:  
Pred. No.: 3.94e-55 Length: 2497  
Score: 511.50 Matches: 106  
Percent Similarity: 62.82% Conservative: 41  
Best Local Similarity: 45.30% Mismatches: 68  
Query Match: 44.48% Indels: 19  
DB: 10 Gaps: 4

US-09-817-199a-2 (1-223) x US-09-834-975-879 (1-2497)

QY 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21  
Db 8 GGCACGCGCC-----CCTCGCGCGCGCGCCCTCCCGCGCTCTCTCCACCGCCT 55  
QY 22 -----CysSerProSerTyrAspLeuThr 29  
Db 56 CCTCTGGCTCCCGGTGTCAGAGGCGCGAGAGAGTGGCGAAGACGTACGATTATCTC 115  
QY 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49  
Db 116 TTCAAGCTCCTGCTGATCGCGGACTCGGGGTAGGCAAGACCTCCCTCTGTTCCGCTTC 175  
QY 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69  
Db 176 TCAGAGAGCGCCTC---AACACCACTTCATCTCCACCATCGAATGATTATTAAT 232  
QY 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89

```
Db 233 AGAAGCATAGAACTAGATGGAAGAAAATAAGCTTCAGATATGGACACACGCGGTGCG 292
Qy 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
Db 293 GAAAGATTCGACAACTCAGACAGCGTACTACAGAGAGCGCATGATGCTGTGTC 352
Qy 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
Db 353 TATGACATCACAAATGAAAATCCTTTGACAATATTAATAAATTGATCAGAAACATTGAA 412
Qy 130 GluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
Db 413 GAGCATGCTCTCCGATGTCGAAAGAAATGATCCTGGGTGGAACAAATGATATGAATGAC 472
Qy 150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
Db 473 AAAGACAAAGTGTCAAAAGAAAGAGGAGAGAGCTAGCAATGACTATGGGATTAATTC 532
Qy 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
Db 533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTACACTTGCACGA 592
Qy 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgasp 208
Db 593 GATATAAATGACAAACTCAACAGAAAATGAATGACAGCAATTCACAGAGCAGGTGGA 652
Qy 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
Db 653 CCAGTGAANAATACAGAAACCGCATCAAGAAGACCAAGTTC 694

RESULT 7
US-09-834-975-885
; Sequence 885, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lilly, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 885
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-885

Alignment Scores:
Pred. No.: 3,94e-55 Length: 2497
Score: 511.50 Matches: 106
Percent Similarity: 62.82% Conservative: 41
Best Local Similarity: 45.30% Mismatches: 68
Query Match: 44.48% Indels: 19
DB: 10 Gaps: 4

US-09-817-199A-2 (1-223) x US-09-834-975-885 (1-2497)

Qy 3 GlyThrProGlyAlaValAlaThrArgaspGlyGluAlaProGluArgSerProPro--- 21
Db 8 GGCACGCC-----CCTCGCGCGCGCGGCCCTCCCGCGCTCTCTCCACCGCCT 55
```

```
Qy 22 -----CysSerProSerTyrAspLeuThr 29
Db 56 CCTCTGGTCCCGCTCAGAGGCGCGAGAGAGATGGCGAAGACGTACGATTATCTC 115
Qy 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
Db 116 TTCAGACGCTCTGCTGATCGGACCTCGGGGTAGCAAGACCTGCTCTGTTCCGCTTC 175
Qy 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
Db 176 TCAGAGGAGCGCTTC---AACACACCTTCATCTCCACCATCGGAATGATTTAAATTC 232
Qy 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
Db 233 AGAAGCATAGAACTAGATGGAAGAAAATAAGCTTCAGATATGGACACACGCGGTGCG 292
Qy 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
Db 293 GAAAGATTCGACAACTCAGACAGCGTACTACAGAGAGCGCATGATGCTGTGTC 352
Qy 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
Db 353 TATGACATCACAAATGAAAATCCTTTGACAATATTAATAAATTGATCAGAAACATTGAA 412
Qy 130 GluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
Db 413 GAGCATGCTCTCCGATGTCGAAAGAAATGATCCTGGGTGGAACAAATGATATGAATGAC 472
Qy 150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
Db 473 AAAGACAAAGTGTCAAAAGAAAGAGGAGAGAGCTAGCAATTCAGCTATGGGATTAATTC 532
Qy 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
Db 533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTACACTTGCACGA 592
Qy 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgasp 208
Db 593 GATATAAATGACAAACTCAACAGAAAATGAATGACAGCAATTCACAGAGCAGGTGGA 652
Qy 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
Db 653 CCAGTGAANAATACAGAAACCGCATCAAGAAGACCAAGTTC 694

RESULT 8
US-09-834-975-894
; Sequence 894, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lilly, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 894
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-894
```

Alignment Scores: 3.94e-55 Length: 2497  
Pred. No.: 511.50 Matches: 106  
Score: 62.82% Conservative: 41  
Percent Similarity: 45.30% Mismatches: 68  
Best Local Similarity: 44.48% Indels: 19  
Query Match: 10 Gaps: 4  
DB:

US-09-817-199a-2 (1-223) x US-09-834-975-894 (1-2497)

```
QY 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro---- 21
|||||
DB 8 GGCACGCC-----CCTCGCGCGCGGCCCTCCCGCTCTCTCCACGGCT 55
QY 22 -----CysSerProSerTyrAspLeuThr 29
|||
DB 56 CTTCTGGCTCCCGGTCAGAGCGCGGAGGAGAGATGGCGAAGACGTACGATTATCTC 115
QY 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPhe 49
|||||
DB 116 TTCAAGCTCTGCTGATCGGAGCTCGGGGTAGGCAAGACCTGCTCTCTCCGCTTC 175
QY 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyLysPheArgAsn 69
|||
DB 176 TCAGAGGAGCGCTTC---AACACCACTTCATCTCCACCATCGGAATTGATTTAAAT 232
QY 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
|||
DB 233 AGAAGATTCGGAACAATAGTAAAGAAATTAAGCTTCAGATATGGACAGCGGGTCTC 292
QY 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
|||
DB 473 AAAAGCAAGTGTCAAAGAAAGAGGGGAGAGAGCTAGCAATTGACTATGGATTAATTC 532
QY 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaLys 189
|||||
DB 533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCAATTTTACACTTGCACGA 592
QY 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
|||
DB 593 GATATATGACAAACTCAACAGAAATGAATGACAGCAATTCAGCAGGAGCGGTGGA 652
QY 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
|||||
DB 653 CCAGTGAATAAACAAGAAACCGATCAAGAAGACCGATTTC 694
```

## RESULT 9

US-09-834-975-896

```
; Sequence 896, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
```

```
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 896
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-896
```

Alignment Scores: 3.94e-55 Length: 2497  
Pred. No.: 511.50 Matches: 106  
Score: 62.82% Conservative: 41  
Percent Similarity: 45.30% Mismatches: 68  
Best Local Similarity: 44.48% Indels: 19  
Query Match: 10 Gaps: 4  
DB:

US-09-817-199a-2 (1-223) x US-09-834-975-896 (1-2497)

```
QY 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro---- 21
|||||
DB 8 GGCACGCC-----CCTCGCGCGCGGCCCTCCCGCTCTCTCCACGGCT 55
QY 22 -----CysSerProSerTyrAspLeuThr 29
|||
DB 56 CTTCTGGCTCCCGGTCAGAGCGCGGAGGAGAGATGGCGAAGACGTACGATTATCTC 115
QY 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPhe 49
|||||
DB 116 TTCAAGCTCTGCTGATCGGAGCTCGGGGTAGGCAAGACCTGCTCTCTCCGCTTC 175
QY 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyLysPheArgAsn 69
|||
DB 176 TCAGAGGAGCGCTTC---AACACCACTTCATCTCCACCATCGGAATTGATTTAAAT 232
QY 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
|||
DB 233 AGAAGATTCGGAACAATAGTAAAGAAATTAAGCTTCAGATATGGACAGCGGGTCTC 292
QY 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
|||
DB 473 AAAAGCAAGTGTCAAAGAAAGAGGGGAGAGAGCTAGCAATTGACTATGGATTAATTC 532
QY 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
|||||
DB 353 TATGACATCAACAATGAAATCCCTTTGACAATATTAAATTTGGATCAGAAACATTGAA 412
QY 130 GluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
|||
DB 413 GAGCATGCTCTCCGATGTGCAAGAAATGATCTCGGTAAACAATGTATATGATGATGAC 472
QY 150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProphe 169
|||
DB 473 AAAAGCAAGTGTCAAAGAAAGAGGGGAGAGAGCTAGCAATTGACTATGGATTAATTC 532
QY 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaLys 189
|||||
DB 533 TTGGAGACAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCAATTTTACACTTGCACGA 592
QY 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
|||
DB 593 GATATATGACAAACTCAACAGAAATGAATGACAGCAATTCAGCAGGAGCGGTGGA 652
QY 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
|||||
DB 653 CCAGTGAATAAACAAGAAACCGATCAAGAAGACCGATTTC 694
```



QY	13	GlyGluAlaProGluArgSer--ProProCysSerProSerTyrAspLeuThrGlyLysV	32
DB	507	GGCGAGAGCCCGCGCGCTCTCCCAATGGCGAAGAAGAGCTACGACCTGCTTTTCAAGC	566
QY	32	alMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspG	52
DB	567	TGCTCTCATCGGGGATTCGGAGTCGGGAAGACCTCGCTCTCTTTTCGTTTTCGGATG	626
QY	52	lyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValV	72
DB	627	ATGCCCTTC---AATACTACCTTTATTTCCACCATAGAGATAGACTTCAAGATCAAAACAG	683
QY	72	alThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgp	92
DB	684	TTCAATTTCAAGCAAAAGACATCAAGCTACAGATATGGGATACAGCAGCGAGGACCGAT	743

```
QY 92 heArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspI 112
Db 744 TCCACACCATCACACACTCTACTACAGAGCGCAATGGGTATCATGCTAGTATATGACA 803
QY 112 leThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrA 132
Db 804 TCACCAATGGTAAAGCTTTGAAACATCAGCAATGGCTTGAGAACATAGATGAGCATG 863
QY 132 laGlnArgaspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgV 152
Db 864 CCAATCAAGATGTGGAAGAATGTTACTAGGAACAAGTGTGATATGGACGACAAAGAG 923
QY 152 aIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluTr 172
Db 924 TTTACCTTAAAGAAAGCAAGACAGATGCAAGGAGCATGTTAGGTATTTTGAGA 983
QY 172 hrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeu 192
Db 984 CTAGTGCAAAGCAATATAACATCGAAAGCGTTCCTCAGTTAGCTGAAGATATCC 1043
QY 192 ystYrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluS 212
Db 1044 TTGCAAGAGACC-----CCTGTAAAGAGAGCCCAACAGTGAATAATGTAGATATCAGCAGTG 1097
QY 212 er-----GlnLysLysArgSerSerCysCys 220
Db 1098 GAGGAGCGGTGACAGCTGGAAGAGCAAAATGCTGC 1132
RESULT 12
US-09-794-257-9
; Sequence 9, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 624
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-9
Alignment Scores:
Pred. No.: 1,11e-54 Length: 624
Score: 501.00 Matches: 98
Percent Similarity: 69.35% Conservative: 40
Best Local Similarity: 49.25% Mismatches: 59
Query Match: 43.57% Indels: 2
DB: 10 Gaps: 2
US-09-817-199a-2 (1-223) x US-09-794-257-9 (1-624)
QY 25 SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCys 44
Db 10 ACGTAGATTATCTCTCAAGCTCTGATCGGCGACTCGGGGTAGCAAGACTGC 69
QY 45 PheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGly 64
Db 70 CTCCTGTCCGCTCTCAGAGGAGCGCTTC---AACACCCTTCATCTCCACCATCGGA 126
QY 65 IleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrp 84
Db 127 ATTGATTTTAAATTAAGACGATAGAACTAGATGGAAGAAATAAGCTTCAGATATGG 186
QY 85 AspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGln 104
```

```
Db 187 GACACAGCGGTGAGAAAGATTCCGAACAATCACGACAGCGTACTACAGAGAGCCATG 246
QY 105 AlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrp 124
Db 247 GCGATTATGCTGGTCTATGACATCACAAATGAAATTCCTTTGACAAATATTAATAATGG 306
QY 125 LeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLys 144
Db 307 ATCAGAAACATTGAAGAGCATGCTCTCCGATGTGCAAGAAATGATCTCTGGGTAAACAA 366
QY 145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164
Db 367 TGTGATATGATGACAAAGACAAAGTGTCAAAAGAGAGGAGGAGGCTAGCAATTGAC 426
QY 165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184
Db 427 TATGGATTAAATCTTGTGAGACAGCGCAAAATCCAGTCCAAATGTAGAAGAGGCATTT 486
QY 185 LeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSer 203
Db 487 TTTACACTTGCAGAGATATAATGACAAAACTCAACAGAAAAAATGAATGACAGCAATTCA 546
QY 204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSerPhe 222
Db 547 GCAGGAGCAGGTGGACCACTGAGTAAATAACAGAAAAACCGATCAAGAAAGACCACTTC 603
RESULT 13
US-09-794-257-7
; Sequence 7, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)...(641)
US-09-794-257-7
Alignment Scores:
Pred. No.: 2,75e-54 Length: 1161
Score: 501.00 Matches: 98
Percent Similarity: 69.35% Conservative: 40
Best Local Similarity: 49.25% Mismatches: 59
Query Match: 43.57% Indels: 2
DB: 10 Gaps: 2
US-09-817-199a-2 (1-223) x US-09-794-257-7 (1-1161)
QY 25 SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCys 44
Db 27 ACGTAGATTATCTCTCAAGCTCTGATCGGCGACTCGGGGTAGCAAGACTGC 86
QY 45 PheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGly 64
Db 87 CTCCTGTCCGCTCTCAGAGGAGCGCTTC---AACACCCTTCATCTCCACCATCGGA 143
QY 65 IleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrp 84
Db 144 ATTGATTTTAAATTAAGACGATAGAACTAGATGGAAGAAATAAGCTTCAGATATGG 203
QY 85 AspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGln 104
```

```
Db 204 GACACAGCGGGTCAGAAAGATTCGGAACAATCACGACAGCGCTACTACAGAGGCCATG 263
Qy 105 AlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrp 124
Db 264 GCATTATGCTGGTCTATGACATCAACAATGAAANAATCCTTGACAAATATTAANAATGG 323
Qy 125 LeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLys 144
Db 324 ATCAGAAACATGGAAGAGCATGCCCTCTCCGATGTCGAAAGAATGATCCTGGTGAACAAA 383
Qy 145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164
Db 384 TGTGATATCAATGACAAAGACAGATGTCGAAAGAATGATCCTGGTGAACAAA 443
Qy 165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184
Db 444 TATGGGTAATTAATCTTGGAGACAGCGCAANAATCCAGTGCAAAATGTAGAAGGCATTT 503
Qy 185 LeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSer 203
Db 504 TTTACACTTGCACGAGATATATGACAAAACCTCAACAGAAAAATGAATGACAGCAATTCA 563
Qy 204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSerPhe 222
Db 564 GCAGGAGCAGGTGGACCAAGTGAATAACAGAAAAACCGATCAAGAAAGACCAAGTTTC 620

RESULT 14
US-09-938-842A-836
; Sequence 836, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 836
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-836

Alignment Scores:
Pred. No.: 1,58e-54 Length: 651
Score: 500.00 Matches: 99
Percent Similarity: 67.61% Conservative: 45
Best Local Similarity: 46.48% Mismatches: 57
Query Match: 43.48% Indels: 12
DB: 9 Gaps: 5

US-09-817-199A-2 (1-223) x US-09-938-842A-836 (1-651)
Qy 19 SerProPro-----CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGly 36
Db 7 GCTCTCTCTGCTGAGAGCTGCTGCTGATTACCTATCAATTAACCTCTGCTGATCGGA 66
Qy 37 AspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer 56
Db 67 GACACGGGTGGTGAAGATTGCCCTCTCTTACGATTCTTCAGATGGCTTCGTTTACCACC 126
Qy 57 GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValIleThrValAspGly 76
```

```
Db 127 ---AGTTTCATTTACAACTATTGGGATTGATTTTAAGATACGGACTATTGAGTCTGATGGG 183
Qy 77 ValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnArgPheArgSerValThr 96
Db 184 AAGACAATTAAGTCGAATCTGGGATCTGCGGACAGACGGGTTCGCCACAATCACA 243
Qy 97 HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSer 116
Db 244 ACTGCGTACTACCGTGGAGCCATGGGATTTTGTGTTGATGATGCTGACTGATGAATCA 303
Qy 117 SerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspVal 136
Db 304 TCTTTCAACAACATCAGGAATTTGGATCCGTAACATTTGAGCAGCATGCTTCTGATAGTGC 363
Qy 137 ValIleMetLeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSer 155
Db 364 AACAGATTCTAGTTGGGAACAACAGCATATGGATGAAGAAAGAAAGACTGTGCCAAA 423
Qy 156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLys 175
Db 424 TCTAAGGSCCAAGCTCTTGCAGATGAATATGATGAATGAAGTTTTCGAGACTAGTGCCAAG 483
Qy 176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195
Db 484 ACTAACTTAAAGCTTGAGGAAGTTTCTTCTCTATTGCTTAAAGACATTAAAGCAAGACTT 543
Qy 196 GlyHis-----GlnAlaAspGluProSerPheGlnIleArgAspTyr----- 209
Db 544 GCAGATACCGATGCAGAGCTGAGCGCGCAACAACAAATCAACCAATCCGACCAAGGT 603
Qy 210 -----ValGluSerGlnLysLysArgSerSerCysCys 220
Db 604 GCGGGAACATCTCAAGCTACTCAGAATCAGCATGTTGC 642

RESULT 15
US-09-938-842A-832
; Sequence 832, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 832
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-832

Alignment Scores:
Pred. No.: 3,98e-54 Length: 609
Score: 496.50 Matches: 93
Percent Similarity: 68.97% Conservative: 47
Best Local Similarity: 45.81% Mismatches: 56
Query Match: 43.17% Indels: 7
DB: 9 Gaps: 3

US-09-817-199A-2 (1-223) x US-09-938-842A-832 (1-609)
Qy 23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42
```

```
Db      4  AATCCTGAATGACATATTGTTCAAGCTTTGCTCATTGGTGATTTCTGGTTGGAAG 63
QY      43  ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr 62
Db      64  TCGTGTGCTTCTTCAAGGTTTGTGATGATTCCTACTGGAT---AGCTACATCAGCACC 120
QY      63  ValGlyIleAspPheArgAsnLysValThrValAspGlyValArgValLysLeuGln 82
Db      121  ATTGGTGTCTGACATTTAAATCCGACATCGAACAAGATGGAAGACCATCAAACTCCAG 180
QY      83  IleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102
Db      181  ATTTGGGATACGGCAGCCAGGAACGTTTCAGAACGATTACTAGCAGTTACTACAGAGGA 240
QY      103  AlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122
Db      241  GCTCATGGAAATTATGTGACTTATGATGTAACAGATCTAGAAAGCTTCAACAACGTCAG 300
QY      123  AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly 142
Db      301  CAATGGCTAAATGAAATCGACCGCTATGCAAGTGAGAATGTTAACAAGCTACTGTTGGG 360
QY      143  AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla 162
Db      361  AACAAGTGTGATCTCATCATCACAAAGTTGTATCCACTGAGACAGCTAAGGCTTTCGCT 420
QY      163  ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182
Db      421  GATGAACCTGGGATCCCATTTCTTGGAAACAAGTGCTAAGAACTCTACCAATGTCGAGAA 480
QY      183  AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGln----- 198
Db      481  GCTTTCATGCCATGACTGCTGCAATCAAGACCAGAAATGGCAGCAACCTCGCAGGAGGA 540
QY      199  AlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSer 218
Db      541  TCTAAGCCACCAACGGTCCAGATCCGAGGACAACCTGTTTAACCAG-----CAATCAGGC 594
QY      219  CysCysSer 221
Db      595  TGCTGCTCC 603
```

Search completed: November 17, 2002, 18:06:57  
Job time : 57 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 17, 2002, 16:40:13 : Search time 254 seconds  
(without alignments)  
1977.148 Million cell updates/sec

Title: US-09-817-199A-2

Perfect score: 1150

Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKRRSCGCSFM 223

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-O=/Cgn2\_1/USPfo\_spool/US09817199/runat\_13112002\_134901\_5153/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCAL\_ALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTENT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817199 -RCGN\_1\_1\_125 @runat\_13112002\_134901\_5153 -NCFU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FPGAPOP=6 -FPGAEXT=7  
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

1: /SID32/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID32/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1150	100.0	1106	22	AAI59066 Human polynucleoti
2	1150	100.0	2323	22	AAI60852 Human polynucleoti
3	1145	99.6	1109	22	AAH75184 Nucleotide sequenc
4	1144	99.5	2623	22	AAH27036 cDNA encoding nove
5	1140	99.1	875	24	ABN83696 Human Ras protein
6	1133	98.5	1316	22	AAH27458 cDNA encoding nove
7	1034	89.9	843	20	AAH86720 DNA sequence encod
8	969	84.3	576	24	ABAO2774 Human degranulatio
9	942	81.9	576	24	ABAO2773 Mouse degranulatio
10	770	67.0	475	23	AAH81431 DNA encoding novel
11	731	63.6	1340	19	AAH65197 Human RAB protein,
12	697	60.6	2001	23	ABL29523 Drosophila melanog
13	535	46.5	964	22	AAH27040 cDNA encoding nove
14	516	44.9	911	21	AAH38429 Arabidopsis thalia
15	515.5	44.8	1129	21	AAH44482 Zea mays DNA fragm
16	515	44.8	1203	21	AAH43685 Zea mays DNA fragm
17	514.5	44.7	2411	24	ABQ54410 Human ovarian anti
18	513.5	44.7	1986	22	ABA09160 Human rab8 homolog
19	511.5	44.5	2497	22	AAH60878 Human cancer agent
20	511.5	44.5	2497	22	AAH60884 Human cancer agent
21	511.5	44.5	2497	22	AAH60893 Human cancer agent
22	511.5	44.5	2497	22	AAH60895 Human cancer agent
23	511.5	44.5	2497	23	ABV25781 Human prostate exp
24	511.5	44.5	2497	23	ABV30037 Human prostate exp
25	508.5	44.2	1540	23	ABL27707 Arabidopsis thalia
26	507	44.1	777	21	AAH42684 Arabidopsis thalia
27	507	44.1	881	24	ABN98761 Arabidopsis thalia
28	507	44.1	1023	21	AAH51491 Arabidopsis thalia
29	507	44.1	1025	21	AAH34347 Arabidopsis thalia
30	507	44.1	3077	22	AAH13912 Human cDNA sequenc
31	506.5	44.0	956	21	AAH96887 Nucleotide sequenc
32	506	44.0	674	23	AAH71453 DNA encoding novel
33	506	44.0	1537	21	AAH16196 Human prostate can
34	506	44.0	3533	22	AAH17889 Human cDNA sequenc
35	505.5	44.0	716	21	AAA40104 Human Rab10 cDNA.
36	505.5	44.0	861	21	AAA40108 Human Rab10 cDNA #
37	503	43.7	959	21	AAH33987 Arabidopsis thalia
38	502	43.7	1101	21	AAH34080 Arabidopsis thalia
39	501	43.6	866	22	AAH04301 Human cDNA clone (
40	501	43.6	1161	22	AAH75182 Nucleotide sequenc
41	498	43.3	888	21	AAH40109 Canine Rab10 cDNA.
42	497.5	43.3	932	21	AAH35200 Arabidopsis thalia
43	490	42.6	2247	23	ABL29661 Drosophila melanog
44	488	42.4	705	21	AAH49208 Arabidopsis thalia
45	483	42.0	666	21	AAH42764 Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAI59066

ID AAI59066 standard; cDNA; 1106 BP.

XX

XX AAI59066;

XX AC

DT 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 1269.

DE

Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.



XX DR WPI; 2001-442253/47.  
DR P-PSDB; AAM41696.  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX PS such as central nervous system injuries -  
XX PS Claim 1; SEQ ID NO 4841; 10078pp; English.  
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX SQ Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Alignment Scores:  
Pred. No.: 6,74e-128 Length: 2323  
Score: 1150.00 Matches: 223  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-817-199A-2 (1-223) x AAI60852 (1-2323)

Qy 1 MetThrGlyThrProGlyValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
Db 28 ATGACGGCAGCCAGGCGCCGTTCACCCAGGATGGCGAGGCCCCCGAGCGCTCCCGC 87  
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
Db 88 CCCTGCAGTCGAGCTACGACCTCACGGGCAAGGTGATGCTTCGGGAGACACAGCGCTC 147  
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
Db 148 GGC AAAACATGTTCTCTGATCCAAATTC AAGACGGGGCTTCCTGTCGGAACCTTCATA 207  
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
Db 208 GCCACCGTCGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGATGAAG 267  
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
Db 268 CTGCAGATCTGGACACCGCTGGCGCAGCAACGGTTCCCGAAGCTCACCCATGCTTATTAC 327  
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
Db 328 AGAGATGCTCAGGCGCTTCCTGCTGTATGATACATCACCAACAATCTCTTTCGACAAAC 387  
Qy 121 IleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
Db 388 ATCAGGGCTGGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447  
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160  
Db 448 CTAGGCAACAAGCGGATATGAGCAGCGCAAGAGGTATCCGCTCCGAGACGGAGAGACC 507  
Qy 161 LeuAlaArgGlyTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180  
Db 508 TTGGCCAGGGAGTACGGTGTTCCTTCCTGAGACACCGCCAGACACTGGCATGAATGTG 567

Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200  
Db 568 GAGTTAGCCTTCTGGCCATCGCCAAAGCACTGAATACCGGGCCGGCATCAGCGGAT 627  
Qy 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220  
Db 628 GAGCCAGCTTCCAGACTCCGAGACTATGTAGAGTCCCAAGAGCGCTCCAGCTGCTGC 687  
Qy 221 SerPheMet 223  
Db 688 TCCCTTCATG 696  
RESULT 3  
AAH75184  
ID AAH75184 standard; DNA; 1109 BP.  
XX AAH75184;  
AC AAH75184;  
XX DT  
XX 13-NOV-2001 (first entry)  
DE Nucleotide sequence of human 32712 G-protein.  
XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;  
KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;  
KW adult respiratory distress syndrome; Goodpasture's syndrome;  
KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;  
KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;  
KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;  
KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;  
KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;  
KW chronic bacterial meningococcalitis; multiple sclerosis;  
KW amyotrophic lateral sclerosis; stroke; Huntington's disease; ss.  
OS Homo sapiens.  
XX FH  
XX Key Location/Qualifiers  
FT CDS 124..699  
FT FT /\*tag= a  
FT FT /product= "G-protein"  
XX PN WO200164887-A2.  
XX PD 07-SEP-2001.  
XX PF 27-FEB-2001; 2001WO-US06292.  
XX PR 29-FEB-2000; 2000US-0185606.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Meyers RA;  
XX WPI; 2001-550182/61.  
DR P-PSDB; AAG67156.  
XX Novel human small G-protein polypeptides and polynucleotides for  
XX treating lung disorders, liver disorders and brain disorders -  
PS Claim 2; Fig 26; 151pp; English.  
XX The present sequence encodes a human G-protein. The specification  
CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The  
CC G-protein polypeptides and polynucleotides are useful as a target for  
CC diagnosis and treatment of G-protein mediated or related disorders,  
CC and for identifying agonists and antagonists for diagnosis and  
CC treatment. They are useful for treating disorders of lung (e.g.  
CC congenital anomalies, pulmonary congestion, oedema, adult respiratory  
CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,  
CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver  
CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,  
CC Wilson's disease, autoimmune hepatitis and hepatic failure), and  
CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute

CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic  
CC bacterial meningoencephalitis, multiple sclerosis, amyotrophic lateral  
CC sclerosis, stroke and Huntington's disease).  
XX  
SQ Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

Alignment Scores:  
Pred. No.: 9,45e-128 Length: 1109  
Score: 1145.00 Matches: 222  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.57% Indels: 0  
DB: 22 Gaps: 0

US-09-817-199A-2 (1-223) x AAH75184 (1-1109)

QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21  
Db 31 ACGGGACACGCCGCGCTTGCACCGGGATGGCGAGGCCCGCCGCGCTCCCGCGCC 90  
QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41  
Db 91 TGCAGTCCGAGCTACACCTCAGGCGCAAGGTGATCTCTGGAGACACAGCGCTCGGC 150  
QY 42 LysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61  
Db 151 AAAACATGTTCTGATCAATCAAGACGGGGCTCTCTGCCGAACCTTCATAGCC 210  
QY 62 ThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81  
Db 211 ACCGTCGGATGACATTCAGAACAGGTGGTGTGTGTGGTGGTGGTGGTGGTGGTGGT 270  
QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101  
Db 271 CAGATCGGACACCGCTGGCGAGAACGGTTCGGAAGCGTCACCCATGCTTATTACAGA 330  
QY 102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121  
Db 331 GATGCTCAGGCGCTGCTGCTGTATGATGATGATGATGATGATGATGATGATGATGAT 390  
QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141  
Db 391 AGGCGCTGCTCACTCAGATTCTATGATGATGATGATGATGATGATGATGATGATGAT 450  
QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161  
Db 451 GGCAACAAGCGGATATGACAGCGAAGAGTATCCGTCGTCGTCGTCGTCGTCGTCGTC 510  
QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181  
Db 511 GCCAGGAGTACGGTGTTCCTTCCTGGAGACCGACGCGCAAGACTGGCATGATGGAG 570  
QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201  
Db 571 TTAGCTTCTGCGCATCGCAAGGAACTGAATATCCGGCGGGCATCAGCGGATGAG 630  
QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221  
Db 631 CCCAGCTTCCAGATCCGAGACTATGATGATGATGATGATGATGATGATGATGATGAT 690  
QY 222 PheMet 223  
Db 691 TTCATG 696

RESULT 4  
AAS27036  
ID AAS27036 standard; cDNA; 2623 BP.  
XX  
AC AAS27036;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE cDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antinflammatory; anti-HIV; antibacterial; rheumatoid arthritis; inflammatory condition;  
KW immune system disorder; infection; hepatitis C; blood disorder;  
KW organ transplant rejection; hyperproliferative disorder; Gaucher's disease;  
KW sickle cell anaemia; Alzheimer's disease; Parkinson's disease;  
KW neurodegenerative disorder; Down syndrome; ischaemia; renal disorder;  
KW chromosomal abnormality; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX  
PN WO200154733-A1.  
XX  
XX 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01312.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 26-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.



```
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0232406.
PR 14-SEP-2000; 2000US-0232423.
PR 21-SEP-2000; 2000US-0234274.
PR 21-SEP-2000; 2000US-0234277.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-SEP-2000; 2000US-0236370.
PR 02-SEP-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-465460/50.
XX P-PSDB; AAU17119.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
XX Claim 1; SEQ ID No 71; 880pp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for
XX diagnosing, preventing and treating diseases including immune system
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
XX transplant rejections and graft versus host disease, infectious diseases
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
XX other blood-related disorders (sickle cell anaemia), myeloproliferative
XX disorders, primary haematopoietic disorders, hyperproliferative
XX disorders (e.g. Gaucher's disease and cancer), neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
XX abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
XX disorders (e.g. glomerulonephritis), cardiovascular disorders
XX (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
XX wound healing, epithelial cell proliferation, endocrine disorders (e.g.
XX Addison's disease), reproductive system disorders, gastrointestinal
XX disorder (inflammatory disorders), liver disorders (cirrhosis),
XX as stimulators of B-cell responsiveness to pathogens, activators of
XX T-cells, to induce higher affinity antibodies, and as a means to induce
XX tumour proliferation in pathologies e.g. acquired immune deficiency
XX syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
XX pathway protein coding sequences and PCR primers of the invention.
XX

Alignment Scores:
Pred. No.: 4.22e-127 Length: 2623
Score: 1144.00 Matches: 222
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 99.48% Indels: 0
DB: 22 Gaps: 0

US-09-817-199A-2 (1-223) x AAS27036 (1-2623)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 12 ATGACGGGACGGCAGCGCGCGTTGCCACCCGGGATGGCGAGGCCGCCGCGCTCCCCG 71
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 72 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTATGCTTCTGGAGACACAGCGGTC 131
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 132 GGCACAAACATGTTTCTCTGATCCAAATTCAGACGGCGGCTTCCTCGCGAACCTTCATA 191
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 192 GCCACCGTCGGCATAGACTTACGACCAAGGTGGTGACTGTGATGGCGGTGAGAGTGAAG 251
```

```
QY 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
DB 252 CTCAGATCTGGACACCGTGGCAGGACGGTCCGAGCGCTCACCCATGCTTATTAC 311
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
DB 312 AGAGATGCTCAGGCTTGCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 371
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
DB 372 ATCAGGCGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGAT 431
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
DB 432 CTAGGCAACAAGCGGATATGACAGCGAAGAGTATCGTTCGCCAAGACGAGAGACC 491
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
DB 492 TTGGCCAGGAGTACGGTGTTCCTTCTGGAGACCAGCGCCCAAGACTGGCATGAATGTG 551
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
DB 552 GAGTTAGCCTTCTGGCCATCGCCAGGAAGTGAATACCGCGCGGCGATCAGGCGGAT 611
QY 201 GluProSerPheClnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
DB 612 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCGAAGAAGCGCTCCAGCTGCTGC 671
QY 221 SerPheMet 223
DB 672 TCCTTCATG 680
RESULT 5
ABN83696
ID ABN83696 standard; cDNA; 875 BP.
XX
AC ABN83696;
XX
XX 27-AUG-2002 (first entry)
XX
DE Human Ras protein 3 (RASP-3) cDNA.
XX
XX Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;
KW immunosuppressive; antiinflammatory; signal transduction; gene;
KW ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 93..668
XX FT /*tag= a
XX FT /product= "RASP-3"
XX
XX US6391580-B1.
XX
XX 21-MAY-2002.
XX
XX 08-MAY-1998; 98US-0075454.
XX
XX 12-DEC-1996; 96US-0766551.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hillman JL, Tang YT, Lal P, Guegler KJ, Corley NC, Patterson C;
XX Batra S, Baughn MR;
XX
XX WPI; 2002-498774/53.
XX P-PSDB; ABB76426.
XX
XX New isolated polynucleotides encoding Ras proteins designated RASP-1
PT and RASP-4, for diagnosing, preventing and treating disorders
PT associated with cell proliferation, particularly cancer and immune
```

```
disorders -
XX
XX Example; Column 57-60; 34pp; English.
XX
CC The present sequence is that of a cDNA clone encoding novel human
CC Ras protein 3 (RASP-3, see ABB76426). Nucleic acids encoding
CC RASP-3 were initially identified in Incyte Clone 1528559 from a
CC mononuclear cell cDNA library (UCMCL5701) using a computer search
CC for amino acid sequence alignments. The present consensus
CC sequence was derived from overlapping and/or extended nucleic
CC acids in Incyte clone 1528559 and shot-gun sequence SARA03135,
CC SAEC10396 and SAEC10855. RASP-3 shows homology to rat Rab26.
CC Northern analysis showed expression of RASP-3 in haematopoietic
CC and immunological cDNA libraries, all associated with inflammation
CC and the immune response. The invention provides 7 novel human Ras
CC proteins (RASP-1 to -7) and polynucleotides, expression vectors,
CC host cells, antibodies, agonists and antagonists. It also provides
CC methods for diagnosing, treating or preventing disorders associated
CC with RASP expression, especially cancer and immune disorders. A
CC fragment of the present sequence, from about nucleotide 92 to about
CC nucleotide 153, is useful as a hybridisation probe.
XX
SQ Sequence 875 BP; 203 A; 247 C; 260 G; 165 T; 0 other;
Alignment Scores:
Pred. No.: 2.7e-127 Length: 875
Score: 1140.00 Matches: 221
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.13% Indels: 0
DB: 24 Gaps: 0
US-09-817-199a-2 (1-223) x ABN83696 (1-875)
QY 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProCys 22
DB 3 GGCACCGCAGGCGCGGTTGCCACCGGGATGGCGAGGCCCCCGAGGCTCCCCGCGCTGC 62
QY 23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42
DB 63 AGTCCGAGCTACGACCTCAGCGGCAGGTGATGCTTCTGGAGACACAGCGGTCGGCAA 122
QY 43 ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr 62
DB 123 ACATGTTCTCTGATCCAAATTCAGAGCGGGCCCTTCCTGTCGGAACCTTCATAGCCACC 182
QY 63 ValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGln 82
DB 183 GTCGGCATACACTTCAGAAACAAGGTGGTGCATGCTGGATGGCGTAGAGTAGAGTCCAG 242
QY 83 IleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102
DB 243 ATCTGGGACACCGCTGGGACAGGAAGCTTCCGAAGCGTCAACCATGCTTATTACAGAGAT 302
QY 103 AlaGlnAlaLeuLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122
DB 303 GCTCAGGCGCTTGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
QY 123 AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly 142
DB 363 GCCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 422
QY 143 AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla 162
DB 423 AACAGCGCGATATGACGACGAGAAAGTATCGTTCGGAAGACGAGAGACCTTGGCC 482
QY 163 ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182
DB 483 AGGAGTACGGTGTTCCTTCTGGAGACGCGCAAGACTGGCATGAATGTGGAGTTA 542
QY 183 AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluPro 202
DB 543 GCCTTCTGGCCATCGCCAGGAAGTGAATACCGGGCCGGGATCAGCGGATGAGCCCC 602
```





PN WO9942586-A2.  
XX 26-AUG-1999.  
XX 23-FEB-1999; 99WO-US03909.  
XX 26-MAY-1998; 98US-0086650.  
PR 23-FEB-1998; 98US-0075534.  
XX (RIGE-) RIGEL PHARM INC.  
XX Anderson D, Fisher J, Huang B, Lorens J, Luo Y;  
PI Shen M;  
XX WPI; 1999-518605/43.  
DR P-PSDB; AAY30132.  
XX New exocytotic proteins useful for diagnosis and treatment of  
PT exocytosis-mediated conditions and in drug screening  
XX Claim 6; Fig 6; 53pp; English.  
XX The present sequence encodes an exocytotic protein designated Exo2.  
CC The protein has an inhibitory effect on exocytosis (i.e. the fusion of  
CC secretory vesicles with the cellular plasma membrane) and can be  
CC administered therapeutically to treat or prevent exocytosis-mediated  
CC disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving  
CC undesirable release of compounds via exocytosis e.g. inflammation  
CC mediated by the release compounds such as histamine. Exo2 is useful  
CC diagnostically and to produce antibodies useful to purify the proteins  
CC and therapeutically to reduce or eliminate the biological activity of  
CC the protein. Exo2 nucleic acids can be used therapeutically to increase  
CC Exo2 activity in cells by known gene therapy techniques. They can also  
CC be used to produce probes or primers to isolate Exo2 proteins from other  
CC organisms, especially humans. The nucleic acids, host cells and proteins  
CC are useful in screening assays to identify binding agents, especially  
CC drug screening assays to identify agonists and antagonists useful  
CC therapeutically to enhance or reduce Exo2 activity.  
XX  
SQ Sequence 843 BP; 195 A; 218 C; 239 G; 175 T; 16 other;

Alignment Scores:  
Pred. No.: 1.43e-114 Length: 843  
Score: 1034.00 Matches: 205  
Percent Similarity: 94.22% Conservative: 7  
Best Local Similarity: 91.11% Mismatches: 11  
Query Match: 89.91% Indels: 2  
DB: 20 Gaps: 0

US-09-817-199A-2 (1-223) x AAX86720 (1-843)

Qy 1 MetThrGlyThrProGlyAla--ValAlaThrArgAspGlyClyAlaProGluArgSerP 20  
Db 26 ATGAATGGCACACAGGAGCTTGCTACCGCTGGGGATGGCGAGCGCCCTGAGCGCTCCC 85  
Qy 20 roProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyV 40  
Db 86 CGCCCTTCAGCCGGAACACTACGATNTCCACGGCAAGTGATGCTTCTTGAGANTCGGCG 145  
Qy 40 aGlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheI 60  
Db 146 TCGGCAAAACCTGTTTCTCTGATCCAAATTCAAAGAGCGGGCTTCTCTCGGAACCTTNA 205  
Qy 60 leAlaThrValClyLeuAspPheArgAsnLysValValThrValAspGlyValArgValL 80  
Db 206 TAGCCACCGTCGGCATACACTACAGATAAAGTGGTACAGTGGATGTCACGGGTGA 265  
Qy 80 ysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrT 100  
Db 266 AGCTTCAGATCTGGGACACTCGACGACAGGAGCGGTCGCGAGTGTGACCCATGCTTAT 325  
Qy 100 yrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspA 120

Db 326 ACCGAGATGCTCAGGCTTTGCTCTGTTGTATGATCATCAACACAGTCTCTTTTGACA 385  
Qy 120 snIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMeTL 140  
Db 386 ACATCAGGCGCTGGCTCACAGAGATTCATGATGCCAGAGGACGCTGGTGATTATGC 445  
Qy 140 euLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluT 160  
Db 446 TTCTAGGCAACAAAGCCGATGTAGCAGCAAGAGGTGATCCGTTCTGAAGATGGAGAGA 505  
Qy 160 hrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAspV 180  
Db 506 CACTGGCCAGGAATATGTTCTTTCATGGACAGCAGTCCCAAGACTGGCATGAAGC 565  
Qy 180 alGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaA 200  
Db 566 TGGAGTTGGCTTCTGGCAATGGCAAGGAACCTGAATACCTGCAGGGAGCGCCTG 625  
Qy 200 spGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysC 220  
Db 626 ATGAGCCAGCTTCCAGATCCGAGACTATGTGGAGTCCCAAGAGAGCGCTCCAGCTGCT 685  
Qy 220 ysSerPheMet 223  
Db 686 GCTCCTTTGTG 696  
RESULT 8  
ABA02774  
ID ABA02774 standard; DNA; 576 BP.  
XX ABA02774;  
AC ABA02774;  
DT 07-FEB-2002 (first entry)  
DE Human degranulation regulator encoding DNA SEQ ID NO 3.  
XX Degranulation; mast cell; human; mouse; antiallergic; ds.  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FT CDS 1..576  
FT /\*tag= a  
FT /product= "degranulation regulator"  
PN WO200179478-A1.  
XX 25-OCT-2001.  
PD 16-APR-2001; 2001WO-JP03268.  
PF 19-APR-2000; 2000JP-0118408.  
PR (DAIN ) DAINIPPON PHARM CO LTD.  
PA Yamada T, Ido M;  
PI WPI; 2002-041335/05.  
XX P-PSDB; AAM53190.  
DR Mast cell degranulation controller for treatment of allergies -  
XX Claim 11; Page 64-65; 85pp; Japanese.  
XX The invention relates to a protein for regulating degranulation of mast  
CC cells (degranulation regulators) and the encoding polynucleotides, with  
CC antiallergic activity, used in the treatment of allergies associated  
CC with degranulation of mast cells.  
XX Sequence 576 BP; 138 A; 147 C; 168 G; 123 T; 0 other;  
SQ Alignment Scores:  
Pred. No.: 5.49e-107 Length: 576

```
Score: 969.00 Matches: 189
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 84.26% Indels: 0
DB: 24 Gaps: 0

US-09-817-199A-2 (1-223) x ABA02774 (1-576)
QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGly 52
Db 1 ATGCTTCTTGGAGACTCGGCGTGGCAAAACATGTTCTCTGATCCAAATCAAGACGGG 60
QY 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
Db 61 GCCTTCTCTCGGAACTTCATAGCACCGTCGGCATAGACTTCAGGAACAAGTGGTG 120
QY 73 ThrValAspGlyValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPhe 92
Db 121 ACTGTGGATGGGTGAGATGAGCTGACAGATCTGGGACCCGCTGGCGAAGACGGTTC 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
Db 181 CGAAGCGTCACCATGCTTATTACAGAGATGCTCAGCGCTTGCTTCTGTGTATGACATC 240
QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
Db 241 ACCAACAATCTCTTTCACACACATCAGGGCTGGCTCCTCAGATTCATGATGATGCC 300
QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerGluArgVal 152
Db 301 CAGAGGAGCTGGTGATCATGCTGTAGGCACACAGCGCGATATGAGCAGCAAGAGTG 360
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 361 ATCCGTTCCGAAGACGAGACCTTGGCCAGGAGTACGGTGTCTCTCTGGAGACC 420
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 421 AGCGCAAGACTGGCATGAATGTGGAGTTAGCCTTCTGGCCATCGCAAGAACTGAAA 480
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 481 TACCGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540

RESULT 9
ABA02773
ID ABA02773 standard; DNA; 576 BP.
XX AC ABA02773;
XX DT 07-FEB-2002 (first entry)
XX DE Mouse degranulation regulator encoding DNA SEQ ID NO 2.
XX KW Degranulation; mast cell; human; mouse; antiallergic; ds.
XX OS Mus musculus.
XX FH Key
XX CDS 1..576
XX FT /*tag= a
XX FT /transl_except= (pos:247..249,aa:Lys)
XX FT /product= "degranulation regulator"
XX PN WO200179478-A1.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-JP03268.
```

```
XX 19-APR-2000; 2000JP-0118408.
XX (DAIN ) DAINIPPON PHARM CO LTD.
XX Yamada T, Ido M;
XX WPI; 2002-041335/05.
XX P-PSDB; AAM52189.
XX Mast cell degranulation controller for treatment of allergies -
PS Claim 10; Page 63-64; 85pp; Japanese.
XX The invention relates to a protein for regulating degranulation of mast
CC cells (degranulation regulators) and the encoding polynucleotides, with
CC anti-allergic activity, used in the treatment of allergies associated
CC with degranulation of mast cells.
XX
SQ Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

Alignment Scores:
Pred. No.: 9 69e-104 Length: 576
Score: 942.00 Matches: 183
Percent Similarity: 98.43% Conservative: 5
Best Local Similarity: 95.81% Mismatches: 3
Query Match: 81.91% Indels: 0
DB: 24 Gaps: 0

US-09-817-199A-2 (1-223) x ABA02773 (1-576)
QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGly 52
Db 1 ATGCTTCTTGGAGACTCGGCGTGGCAAAACCTGTTCTCTGATCCAAATCAAGACGGG 60
QY 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
Db 61 GCCTTCTCTCGGAACTTCATAGCACCGTCGGCATAGACTTCAGGAATAAAGTGGTG 120
QY 73 ThrValAspGlyValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPhe 92
Db 121 ACAGTGGATGGTGGCCAGGCTGAAGCTTCAGATCTGGACACTGCAGACAGGAGCCCTTT 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
Db 181 CGCAGTGTGACCATGCTTATTACCGAGATGCTCAGGCTTGTCTCTGTGTATGACATC 240
QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
Db 241 ACCAACCAGTCTCTTTTGACACATCAGGGCTGGCTCAGAGATTCATGATGATGCC 300
QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerGluArgVal 152
Db 301 CAGAGAGACGTGGTGATTCCTTCTAGGCAACAAGCCGCTAGAGCAGCAAGAGGTG 360
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 361 ATCCGTTCTGAAGATGGAGAGACTGCGCAGGGAATATGGTGTCTCTTTCATGGAGACC 420
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 421 AGTGCCAAGACTGGCATGAACGTGGAGTGGCTTCTTGCAATTCGCAAGAACTGAAA 480
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 481 TACCGTGGCAGGAGGAGCCGCTGTATGAGCCAGCTTCCAGATCCGAGACTATGTGGAGTCC 540
QY 213 GlnLysLysArgSerSerCysCysSerPheMet 223
Db 541 CAGAAGAAGCGCTCCAGCTGCTCTCTTTGTG 573

RESULT 10
AAS81431
```

AA881431 standard; cDNA; 475 BP.  
AA881431;  
13-FEB-2002 (first entry)  
DNA encoding novel human diagnostic protein #17235.  
Human; chromosome mapping; gene mapping; gene therapy; forensic;  
food supplement; medical imaging; diagnostic; genetic disorder; ss.  
Homo sapiens.  
WO200175067-A2.  
11-OCT-2001.  
30-MAR-2001; 2001WO-US08631.  
31-MAR-2000; 2000US-0540217.  
23-AUG-2000; 2000US-0649167.  
(HYSE-) HYSEQ INC.  
Drmanac RT, Liu C, Tang YT;  
WPI; 2001-639362/73.  
P-PSDB; ABG17244.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -  
Claim 1; SEQ ID No 17235; 103pp; English.  
The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags  
for identifying expressed genes. (II) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. AA881431-17235 represent novel human  
diagnostic coding sequences of the invention.  
Note: The sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;  
Alignment Scores:  
Pred. No.: 3.58e-83 Length: 475  
Score: 770.00 Matches: 148  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.33% Mismatches: 0  
Query Match: 66.96% Indels: 0  
DB: 23 Gaps: 0  
US-09-817-199A-2 (1-223) x AA881431 (1-475)  
Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
Db 28 ATGACGGGACGACGAGGCGCGTTCACCGGGATGGCGAGGCCCGCGGCTCCCGC 87

Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
Db 88 CCCTGCAGTCCGAGCTACGACCTCAGGCAAGGTGATGCTTCTGGGACACAGCGGTC 147  
Qy 41 GlyLysThrCysPheLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
Db 148 GCAAAACATGTTCTCTGATCCAAATCAAGACGGGGCTTCTGTCGGAACCTTCATA 207  
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
Db 208 GCCACCGCGGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAAGTGAAG 267  
Qy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
Db 268 CTGCAGATCTGGACACCGCTGGCGAGGACGGTCCGAAGCGTCACCCATCTATTATAC 327  
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
Db 328 AGAGATGCTCAGGCGCTTGTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 387  
Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
Db 388 ATCAGGGCTGGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
Qy 141 LeuGlyAsnLysAlaAspMetSerSer 149  
Db 448 CTAGGCAACAAGCGCGATATGAGCAGC 474  
RESULT 11  
AAV65197  
ID AAV65197 standard; DNA; 1340 BP.  
XX AAV65197;  
XX  
XX 17-DEC-1998 (first entry)  
XX Human RAB protein, SRAB, coding sequence.  
XX  
XX Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;  
XX Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;  
XX viral infection; therapy; autoimmune sialosis; cystic fibrosis;  
XX diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;  
XX hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;  
XX polycystic renal disease; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 43..615  
XX FT /\*tag= a  
XX FT /product= "SRAB"  
XX  
XX WO9842839-A1.  
XX  
XX PD 01-OCT-1998.  
XX  
XX PF 25-MAR-1998; 98WO-US05871.  
XX  
XX PR 26-MAR-1997; 97US-0824873.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Guegler KJ, Hillman JL;  
XX  
XX WPI; 1998-543282/46.  
XX DR P-PSDB; AAW80747.  
XX  
XX New human Rab protein, SRAB - useful for treating disorders  
XX associated with SRAB expression, including vesicle trafficking,  
XX viral infection, and cancer  
XX Claim 5; Fig 1; 31pp; English.





Db	1477	TATGTGCGGAGTATTTCTCAGCACAGTGTGGCAATGATTTTAGGAACAAAAGTGGTGTC	1536
Qy	74	ValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArg	93
Db	1537	GTCGATGAACGGCGTCAAGCTGCAATCTGGACACAGCTGGTCAGGAGCGATTCCGG	1596
Qy	94	SerValThrHisAlaTyrrArgAspAlaGlnAlaLeuLeuLeuLeuTyrrAspileThr	113
Db	1597	AGCGTTACCCAGCGCTATTATCGGACGCGCAGCTCTACTGCTGCTACGACGTGACC	1656
Qy	114	AsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrrAlaGln	133
Db	1657	AACAGACCACTATGACAACATTCGCGCCTGGCTGGCGGAGATCCGGGATACGCGAG	1716
Qy	134	ArgAspValValIleMetLeuGlyAsnLysAlaAspMetSer---SerGluArgVal	152
Db	1717	GAGGACGTGGTCATCGTTTTATAGGCACAAAGGCCGACTGCAGCGGAGCGGCGGAG	1776
Qy	153	IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrrGlyValProPheLeuGluThr	172
Db	1777	GTGAAGCGGAGGATGGGAGCGTTTGGGCGGGAGCACACAGTGCCTTCATGAGACC	1836
Qy	173	SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys	192
Db	1837	TCGGCCAAAGCGGACTCAATGTGGAGCTGTCTTCACAGCGGTGCCAGGCAACTAAAG	1896
Qy	193	TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrrValGluSer	212
Db	1897	AGTCGGCGCTAGGACACGCGGATGTGGAAGTTCAATGTGCATGATTTTGTGCGTGAC	1956
Qy	213	GlnLysLysArgSerSerCysCysSer 221	
Db	1957	AATACAAAGCGCGCTCGGTTTGCGCC 1983	
RESULT 13			
ID	AAS27040	standard; cdna; 964 BP.	
XX	XX		
AC	AAS27040;		
XX	XX		
DT	07-NOV-2001	(first entry)	
XX	XX		
DE	XX		
XX	XX		
KW	XX	cdna encoding novel signal transduction pathway protein, Seq ID 75.	
KW	XX	Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;	
KW	XX	antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;	
KW	XX	immune system disorder; rheumatoid arthritis; inflammatory condition;	
KW	XX	organ transplant rejection; infection; hepatitis C; blood disorder;	
KW	XX	sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;	
KW	XX	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;	
KW	XX	chromosomal abnormality; Down syndrome; ischaemia; renal disorder;	
KW	XX	cardiovascular; respiratory; wound healing; endocrine; Addison's disease;	
KW	XX	reproductive system; gastrointestinal; liver disorder; AIDS; ss;	
KW	XX	acquired immune deficiency syndrome.	
OS	XX	Homo sapiens.	
XX	XX		
PN	XX	WO200154733-A1.	
XX	XX		
PD	XX	02-AUG-2001.	
XX	XX		
PF	XX	17-JAN-2001; 2001WO-US01312.	
XX	XX		
PR	31-JAN-2000;	2000US-0179065.	
PR	04-FEB-2000;	2000US-0180628.	
PR	24-FEB-2000;	2000US-0184664.	
PR	02-MAR-2000;	2000US-0186350.	
PR	16-MAR-2000;	2000US-0189874.	
PR	17-MAR-2000;	2000US-0190076.	
PR	18-APR-2000;	2000US-0198123.	
PR	19-MAY-2000;	2000US-0205515.	
PR	07-JUN-2000;	2000US-0209467.	
PR	28-JUN-2000;	2000US-0214886.	
PR	30-JUN-2000;	2000US-0215135.	
PR	07-JUL-2000;	2000US-0216647.	
PR	07-JUL-2000;	2000US-0216880.	
PR	11-JUL-2000;	2000US-0217487.	
PR	11-JUL-2000;	2000US-0217496.	
PR	14-JUL-2000;	2000US-0218290.	
PR	26-JUL-2000;	2000US-0220963.	
PR	26-JUL-2000;	2000US-0220964.	
PR	14-AUG-2000;	2000US-0224518.	
PR	14-AUG-2000;	2000US-0224519.	
PR	14-AUG-2000;	2000US-0225213.	
PR	14-AUG-2000;	2000US-0225214.	
PR	14-AUG-2000;	2000US-0225266.	
PR	14-AUG-2000;	2000US-0225267.	
PR	14-AUG-2000;	2000US-0225268.	
PR	14-AUG-2000;	2000US-0225270.	
PR	14-AUG-2000;	2000US-0225447.	
PR	14-AUG-2000;	2000US-0225757.	
PR	14-AUG-2000;	2000US-0225758.	
PR	14-AUG-2000;	2000US-0225759.	
PR	18-AUG-2000;	2000US-0226279.	
PR	22-AUG-2000;	2000US-0226681.	
PR	22-AUG-2000;	2000US-0226688.	
PR	22-AUG-2000;	2000US-0227182.	
PR	23-AUG-2000;	2000US-0227009.	
PR	30-AUG-2000;	2000US-0228924.	
PR	01-SEP-2000;	2000US-0229287.	
PR	01-SEP-2000;	2000US-0229343.	
PR	01-SEP-2000;	2000US-0229344.	
PR	01-SEP-2000;	2000US-0229345.	
PR	05-SEP-2000;	2000US-0229509.	
PR	05-SEP-2000;	2000US-0229513.	
PR	06-SEP-2000;	2000US-0230437.	
PR	06-SEP-2000;	2000US-0230438.	
PR	08-SEP-2000;	2000US-0231242.	
PR	08-SEP-2000;	2000US-0231243.	
PR	08-SEP-2000;	2000US-0231244.	
PR	08-SEP-2000;	2000US-0231413.	
PR	08-SEP-2000;	2000US-0231414.	
PR	08-SEP-2000;	2000US-0232080.	
PR	08-SEP-2000;	2000US-0232081.	
PR	12-SEP-2000;	2000US-0231968.	
PR	14-SEP-2000;	2000US-0232397.	
PR	14-SEP-2000;	2000US-0232398.	
PR	14-SEP-2000;	2000US-0232399.	
PR	14-SEP-2000;	2000US-0232400.	
PR	14-SEP-2000;	2000US-0232401.	
PR	14-SEP-2000;	2000US-0233063.	
PR	14-SEP-2000;	2000US-0233064.	
PR	14-SEP-2000;	2000US-0233065.	
PR	21-SEP-2000;	2000US-0234223.	
PR	21-SEP-2000;	2000US-0234224.	
PR	25-SEP-2000;	2000US-0234997.	
PR	25-SEP-2000;	2000US-0234998.	
PR	27-SEP-2000;	2000US-0235484.	
PR	27-SEP-2000;	2000US-0235836.	
PR	29-SEP-2000;	2000US-0236327.	
PR	29-SEP-2000;	2000US-0236367.	
PR	29-SEP-2000;	2000US-0236368.	
PR	29-SEP-2000;	2000US-0236369.	
PR	29-SEP-2000;	2000US-0236370.	
PR	02-OCT-2000;	2000US-0236802.	
PR	02-OCT-2000;	2000US-0237037.	
PR	02-OCT-2000;	2000US-0237038.	
PR	02-OCT-2000;	2000US-0237039.	
PR	13-OCT-2000;	2000US-0237040.	
PR	13-OCT-2000;	2000US-0239935.	
PR	13-OCT-2000;	2000US-0239937.	
PR	20-OCT-2000;	2000US-0240960.	
PR	20-OCT-2000;	2000US-0241221.	
PR	20-OCT-2000;	2000US-0241785.	
PR	20-OCT-2000;	2000US-0241786.	



XX	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055.
PD		PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144884.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	19-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	20-MAY-1999;	PR	28-JUL-1999;	99US-0145911.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	03-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151338.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	PR	20-SEP-1999;	99US-0154779.

PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
Pred. No.: 3,18e-52 Length: 911  
Score: 516.00 Matches: 107  
Percent Similarity: 66.52% Conservative: 48  
Best Local Similarity: 45.92% Mismatches: 58  
Query Match: 44.87% Indels: 20  
DB: 21 Gaps: 6

US-09-817-199a-2 (1-223) x AAC38429 (1-911)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20  
Db 28 CTAACATCTTCTCGATGCGGTT-----GCGCGCGCAAGAGCT--- 66  
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40  
Db 67 -----CGTTACAGACTATGATTATCTCATCAAGCTTCTCTCATCGGTGATAGCGGTG 120  
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60  
Db 121 GGAAGAGTCTGTTACTGCGGTTTCGATGATACTTCCACCAG---AGTTTCATT 177  
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80  
Db 178 ACTACCATTTGAATCGACTTCAAGATAAGACAGCTCGAAGCTTGATGGAAGCGTATCAAA 237  
Qy 81 LeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100  
Db 238 TTGCAGATATGGACATCGCTGSGACAAGCGTTTATAGACTATAACACAGCATATTAC 297  
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120  
Db 298 AGAGGAGCGATGGGTATATTACTCGTCTATGATATACAGATGAGTCATCATCAACAAAC 357

Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140  
Db 358 ATTAGGAACTGGATGAAAAACATTGAGCAGCATGCGTCCGATAGTGTCAACAAGATATTG 417  
Qy 141 LeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSerGluAspGlyGlu 159  
Db 418 GTCGGTAAAGCGGACATGAGCAAGAAAGGCTGTCCCAACATCAAAAGGACAA 477  
Qy 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179  
Db 478 GCTCTTGCTGATGATGGAATCAATCTTTGAGAGGAGTGCACAAACAAACACAGAAT 537  
Qy 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg-----AlaGly 196  
Db 538 GTCGAGCAGGTTTCTTCTATCGCTAAAGACATAAAACAAAGACTCACAGAAAGCGAT 597  
Qy 197 HisGlnAlaAspGluProSerPheGlnIle-----ArgAspTyrVal 210  
Db 598 ACAAAAGCGGAGCCCAAGGGATCAAGATCACATAACAAGATGCTAACAAAGCCTCATCG 657  
Qy 211 GluSerGlnLysLysArgSerSerCysCysSerPheMet 223  
Db 658 TCCTTCTACAAATGAGAAATCAGCTTGTCTGCACTTATGTT 696  
RESULT 15  
AAC44482  
ID AAC44482 standard; DNA; 1129 BP.  
AC AAC44482;  
XX  
XX 18-OCT-2000 (first entry)  
DE Zea mays DNA fragment SEQ ID NO: 42986.  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic;  
KW pathway; promoter; termination sequence; corn; ss.  
OS Zea mays subsp. mays.  
PN EP1033405-A2.  
XX PD 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144633.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4,95e-52 Length: 1129
Score: 515.50 Matches: 104
Percent Similarity: 66.38% Conservative: 50
Best Local Similarity: 44.83% Mismatches: 61
Query Match: 44.83% Indels: 17
DB: 21 Gaps: 6

US-09-817-199a-2 (1-223) x AAC44482 (1-1129)
QY 5 ProGlyAlaValAlaThrArgAspGlyGluAlaProGluArg----- 18
Db 127 CCGGGACCGCGGGTCCGCGCGGGTAGGAGGAGGAGGAGGCGGATGCGG 186
QY 19 SerProPro-----CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGly 36
Db 187 GCGCGCGCGCGAGGCGCGCGGCGGACTACGACTACCTTATCAAGCTGCTCTCATTTGA 246
QY 37 AspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSer 56
Db 247 GACAGCGGTGTGGCAAGAGTCCCTCCCTGTTGCGGTTCTCTGTTGTTCCCTTCACTACA 306
QY 57 GlyThrPheLeuAlaThrValGlyLeuAspPheArgAsnLysValValThrValAspGly 76
Db 307 ---AGCTTTATTACCACAAATGGTATGACTTTAAGATAGCAACGATAGAAATGGATGGC 363
QY 77 ValArgValLysLeuGlnLeuIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThr 96
Db 364 AAGCGTATAAAGTACAGATTTGGGATACGCGCGGCGGCAAGACGCTTCCGGAATATTACC 423
QY 97 HisAlaTyrTrpArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSer 116
Db 424 ACCCGCTACTACCGAGGAGCATGCGGCACTCTGCTGTTTATGATGTCACCGAATGATCT 483
QY 117 SerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspVal 136
Db 484 TCITTTAACAACTCCGAACTGGATTTCGGCAACATTGAACACACACGCGCTCTGATAATGTC 543
QY 137 ValIleMetLeuLeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSer 155
Db 544 AACAAAGATTTTGGTTGGCAACAGCGCGGATGAGCAAGCAAGAGGGGCTGTACCTACT 603
QY 156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLys 175
Db 604 GCAAAAGGACAAAGCATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
QY 176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195
Db 664 ACAACCTGATGTGGAGCAAGTCTTTTCTCCATCCGACGTGACATCAAGCAGAGGCTT 723
QY 196 GlyHisGlnAlaAspGluPro-----SerPheGlnIleArgAspTyrValGluSer 212
Db 724 GCGGAGACTGATTCAAAGCCAGAGGACAGGCAATCAAAATTAACAAAGCCAGACCGGCT 783
QY 213 GlnLys-----LysArgSerSerCysCys 220
Db 784 TCCGAGGACCCAGCTGCCAGCGATCTGCTGTGTC 819
```

Search completed: November 17, 2002, 16:48:15  
Job time : 258 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 16:39:28 ; Search time 20 Seconds  
(without alignments)  
167.929 Million cell updates/sec

Title: US-09-817-199A-2

Perfect score: 1150

Sequence: 1 MTGTPGAVATRDGEAPERSP.....FOIRDYVESQKRSSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB\_PEP.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB\_PEP.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB\_PEP.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB\_PEP.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB\_PEP.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB\_PEP.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1150	100.0	223	10	US-09-817-199A-2
2	1081	94.0	223	10	US-09-817-199A-4
3	977	85.0	191	10	US-09-794-257-14
4	977	85.0	191	12	US-10-051-986-3
5	547	47.6	106	10	US-09-867-550-1812
6	501	43.6	207	10	US-09-794-257-8
7	500	43.5	190	10	US-09-822-860-5
8	498	43.3	218	10	US-09-925-300-1571
9	497.5	43.3	162	10	US-09-834-765-766
10	483	42.0	245	10	US-09-925-302-534
11	472	41.0	201	10	US-09-967-736-8
12	468	40.7	201	10	US-09-967-736-3
13	466	40.5	198	10	US-09-794-257-16
14	466	40.5	198	10	US-09-945-173-5
15	466	40.5	198	10	US-09-972-529-4
16	458.5	39.9	222	10	US-09-820-003A-4
17	446.5	39.0	212	10	US-09-350-874-67
18	429	37.3	212	10	US-09-817-198A-2
19	427.5	37.2	212	10	US-09-817-198A-4

20	416	36.2	218	10	US-09-817-198A-5
21	413.5	36.0	312	10	US-09-925-302-783
22	386	33.6	832	10	US-09-834-765-2
23	373.5	32.5	213	10	US-09-794-257-5
24	365	31.7	208	9	US-10-108-605-45
25	364.5	31.7	624	10	US-09-834-765-5
26	364.5	31.7	625	10	US-09-834-765-762
27	360.5	31.3	168	10	US-09-834-765-765
28	360	31.3	239	10	US-09-925-301-1077
29	359	31.2	217	10	US-09-988-974-3
30	358.5	31.2	213	10	US-09-988-974-8
31	352.5	30.7	201	10	US-09-925-300-1364
32	350	30.4	217	10	US-09-822-860-2
33	349.5	30.4	161	10	US-09-834-765-763
34	344.5	30.0	211	12	US-10-051-986-6
35	339	29.5	216	10	US-09-945-173-10
36	319.5	27.8	157	10	US-09-834-765-764
37	314.5	27.3	173	10	US-09-820-003A-2
38	313.5	27.3	206	10	US-09-828-310-13
39	305.5	26.6	201	10	US-09-988-974-9
40	302.5	26.3	201	10	US-09-988-974-5
41	302.5	26.3	209	10	US-09-864-761-42996
42	299.5	26.0	208	10	US-09-925-302-629
43	294.5	25.6	137	10	US-09-817-182-6
44	284.5	24.7	259	12	US-10-051-986-1
45	284.5	24.7	260	12	US-10-051-986-4

ALIGNMENTS

RESULT 1

US-09-817-199A-2  
; Sequence 2, Application US/09817199A  
; Patent No. US20(2014)2380A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001187  
; CURRENT APPLICATION NUMBER: US/09/817,199A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 223  
; TYPE: PRT  
; ORGANISM: Human  
US-09-817-199A-2

Query Match	100.0%	Score	1150;	DB	10;	Length	223;
Best Local Similarity	100.0%	Pred. No.	1.5e-116;				
Matches	223;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MTGTPGAVATRDGEAPERSPPCPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI	60				
Db	1	MTGTPGAVATRDGEAPERSPPCPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI	60				
QY	61	ATVGDIDFNKVVTVDGVVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDN	120				
Db	61	ATVGDIDFNKVVTVDGVVRVKLQIWDTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDN	120				
QY	121	IRAWLTIHEYAQRDVIMLLGNKADMSSEVRISDEGTETLAREYGVPLETSAKTGMNV	180				
Db	121	IRAWLTIHEYAQRDVIMLLGNKADMSSEVRISDEGTETLAREYGVPLETSAKTGMNV	180				
QY	181	ELAFIAIAKELKYRAGHQADEPSFOIRDYVESQKRSSCCSFM	223				
Db	181	ELAFIAIAKELKYRAGHQADEPSFOIRDYVESQKRSSCCSFM	223				

RESULT 2

US-09-817-199a-4  
; Sequence 4, Application US/09817199a  
; Patent No. US20020142380A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001187  
; CURRENT APPLICATION NUMBER: US/09/817,199a  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 223  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-817-199a-4

Query Match 94.0%; Score 1081; DB 10; Length 223;  
Best Local Similarity 93.7%; Pred. No. 4.2e-109;  
Matches 209; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MCTPGCAVTRDGEAPERSPPSPDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60  
Db 1 MCTPGCAATAGDGEAPERSPPSPDLTGKVMLLGDSGVGKTCFLIQKDGAFLSGTFI 60  
QY 61 ATVGIDFRNKVTVGVRVKLQIWDTAGOERPRSVTHAYYRDAQALLLLYDITNKSSFDN 120  
Db 61 ATVGIDFRNKVTVGVRVKLQIWDTAGOERFRSVTHAYYRDAQALLLLYDITNKSSFDN 120  
QY 121 IRALWTEIHEYAQRDVVIMLLGNKADMSSEVIRSEDGETLAREYGVFPFLETSKATGMNV 180  
Db 121 IRALWTEIHEYAQRDVVIMLLGNKADVSSEVIRSEDGETLAREYGVFPFLETSKATGMNV 180  
QY 181 ELAFLAIKELKYRAGHQADEPFSQIRDYVESQKRRSSCCSFM 223  
Db 181 ELAFLAIKELKYRAGQPDEPFSQIRDYVESQKRRSSCCSFM 223

RESULT 3  
US-09-794-257-14  
; Sequence 14, Application US/09794257  
; Patent No. US20020009804A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1  
; FILE REFERENCE: 35800/209285  
; CURRENT APPLICATION NUMBER: US/09/794,257  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/185,606  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-794-257-14

Query Match 85.0%; Score 977; DB 10; Length 191;  
Best Local Similarity 100.0%; Pred. No. 5.5e-98;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVGVRVKLIQIWDTAGQERF 92  
Db 1 MLLGDTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVGVRVKLIQIWDTAGQERF 60  
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRALWTEIHEYAQRDVVIMLLGNKADMSSEV 152  
Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRALWTEIHEYAQRDVVIMLLGNKADMSSEV 120

QY 153 IRSEDGETLAREYGVFPFLETSKATGMNVFLAFLAIKELKYRAGHQADEPFSQIRDYVES 212  
Db 121 IRSEDGETLAREYGVFPFLETSKATGMNVFLAFLAIKELKYRAGHQADEPFSQIRDYVES 180  
QY 213 QKRRSSCCSFM 223  
Db 181 QKRRSSCCSFM 191  
RESULT 4  
US-10-051-986-3  
; Sequence 3, Application US/10051986  
; Patent No. US20020146770A1  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; Tang, Y. Tom  
; Lal, Preeti  
; Guegler, Karl J.  
; Corley, Neil C.  
; Patterson, Chandra  
; Batra, Sajeev  
; Baughn, Mariah R.  
; TITLE OF INVENTION: RAS PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/051,986  
; FILING DATE: 15-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/766,551  
; FILING DATE: DECEMBER 12, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cerrone, Michael C.  
; REGISTRATION NUMBER: 39,132  
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 191 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: UCMCL5T01  
; CLONE: 1528559  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3 :  
US-10-051-986-3

Query Match 85.0%; Score 977; DB 12; Length 191;  
Best Local Similarity 100.0%; Pred. No. 5.5e-98;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVGVRVKLIQIWDTAGQERF 92  
Db 1 MLLGDTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVGVRVKLIQIWDTAGQERF 60  
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRALWTEIHEYAQRDVVIMLLGNKADMSSEV 152



Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSERV 120  
Qy 153 IRSEDETLAREYGVFPFLETSAKTGMNVELAFALAKELKYRAGHOADEPSFOIRDYVES 212  
Db 121 IRSEDETLAREYGVFPFLETSAKTGMNVELAFALAKELKYRAGHOADEPSFOIRDYVES 180  
Qy 213 QKRSSCCSFM 223  
Db 181 QKRSSCCSFM 191

RESULT 5  
US-09-867-550-1812  
; Sequence 1812, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1812  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-1812

Query Match 47.6%; Score 547; DB 10; Length 106;  
Best Local Similarity 100.0%; Pred. No. 5.5e-52;  
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 33 MLLGDTGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLQIWDTAGQERF 92  
Db 1 MLLGDTGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLQIWDTAGQERF 60  
Qy 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVI 138  
Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVI 106

RESULT 6  
US-09-794-257-8  
; Sequence 8, Application US/09794257  
; Patent No. US2002009804A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1el  
; FILE REFERENCE: 35800/209285  
; CURRENT APPLICATION NUMBER: US/09/794,257  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/185,606  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-794-257-8

Query Match 43.6%; Score 501; DB 10; Length 207;  
Best Local Similarity 49.2%; Pred. No. 1.3e-46;  
Matches 98; Conservative 40; Mismatches 59; Indels 2; Gaps 2;

Qy 25 SYDLGCKVMLGDTGVGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLQIW 84  
Db 4 TYDLFKLLIGDSGVGKTCFLFRSEDAF-NTTISTIGIDFKIRTIELDGKKIKLQIW 62  
Qy 85 DTAGQERFSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNK 144  
Db 63 DTAGQERFRTITAYYRGAMGIMLYDITNEKSFNKNWIRNIEEHASDVVERMILGNK 122  
Qy 145 ADMSEFVIRSEDETLAREYGVFPFLETSAKTGMNVELAFALAKELKYRAGHOA-DEPS 203  
Db 123 CDMDKQVSKERGEKLAIDYGIKLETSAKSSANVEEAFFTLARDINTKLNRKMNDSNS 182  
Qy 204 FOIRDYVESQKRSCCSF 222  
Db 183 AGAGGPVKITENRKNKTSF 201  
RESULT 7  
US-09-822-860-5  
; Sequence 5, Application US/09822860  
; Patent No. US20020146795A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHU, Shiaoqing et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; FILE REFERENCE: CLO01214  
; CURRENT APPLICATION NUMBER: US/09/822,860  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Discopyge ommata  
US-09-822-860-5  
Query Match 43.5%; Score 500; DB 10; Length 190;  
Best Local Similarity 50.3%; Pred. No. 1.5e-46;  
Matches 96; Conservative 42; Mismatches 51; Indels 2; Gaps 2;  
Qy 27 DLTGKVMMLGDTGVGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLQIWD 86  
Db 1 DYLFKLLIGDSGVGKTCFLFRSEDAF-NTTISTIGIDFKIRTVELDGKKIKLQIWD 59  
Qy 87 AQQERFSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKAD 146  
Db 60 AQQERFRTITAYYRGAMGIMKYDITNEKSFNKNWIRNIEEHASDVVERMILGNKCD 119  
Qy 147 MSSERVIRSEDETLAREYGVFPFLETSAKTGMNVELAFALAKELKYRAGHOADEPSFQ- 205  
Db 120 MNEKQVSKERGEKLAIDYGIKLETSAKSSINVEEAFFTLARDINTKLNRKMNENSLQ 179  
Qy 206 IRDYVESQKKR 216  
Db 180 AVDKLKSPPKK 190  
RESULT 8  
US-09-925-300-1571  
; Sequence 1571, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270



NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 57006  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-967-736-8

Query Match 41.0%; Score 472; DB 10; Length 201;  
Best Local Similarity 43.1%; Pred. No. 1.7e-43;  
Matches 87; Conservative 51; Mismatches 58; Indels 6; Gaps 3;

QY 23 SPSYDLTKVMLLGDGTGKTCFLIQFDGAFSLGCTFIATVGDIFRNKVVTVVDGVRVKIQ 82  
Db 2 NPEYDLFKLLIGDSGVGKSCLLLRFAADTY-TESYSTIGVDKIRTIELDGKTIKIQ 60

QY 83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFFDNIRAWLTIHEHYAQRDVVIMLLG 142  
Db 61 IWDTAGQERFRVTSYRGAGHIIIVYDVTDQESYANVKQLQEIYDRIYASENVNKLIVG 120

QY 143 NKADSSRVRIRSEGETLAREYGVFPFLETSAKTGMNVELAFALAIKELKYRAGHOA--- 199  
Db 121 NKSDLTKKVVNDTTAKEFADSLGVFPFLETSAKNATNVEQAPMTMAAEIKKRMGPGAASG 180

QY 200 -DEPSFQIRDYVESQKRSSCC 220  
Db 181 GERPNLKI-DSTPVKASGGCC 201

RESULT 12  
US-09-967-736-3  
Sequence 3, Application US/09967736  
Patent No. US20020103340A1  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
Lal, Preeti  
Corley, Neil C.  
Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/967,736  
FILING DATE: 28-Sep-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/154,602  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LivTUT04  
CLONE: 2514506  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-967-736-3

Query Match 40.7%; Score 468; DB 10; Length 201;  
Best Local Similarity 42.1%; Pred. No. 4.6e-43;  
Matches 85; Conservative 53; Mismatches 58; Indels 6; Gaps 3;

QY 23 SPSYDLTKVMLLGDGTGKTCFLIQFDGAFSLGCTFIATVGDIFRNKVVTVVDGVRVKIQ 82  
Db 2 NPEYDLFKLLIGDSGVGKSCLLLRFAADTY-TESYSTIGVDKIRTIELDGKTIKIQ 60

QY 83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFFDNIRAWLTIHEHYAQRDVVIMLLG 142  
Db 61 IWDTAGQERFRVTSYRGAGHIIIVYDVTDQESYANVKQLQEIYDRIYASENVNKLIVG 120

QY 143 NKADSSRVRIRSEGETLAREYGVFPFLETSAKTGMNVELAFALAIKELKYRAGHOA--- 199  
Db 121 NKSDLTKKVVNDTTAKEFADSLGVFPFLETSAKNATNVEQAPMTMAAEIKKRMGPGAASG 180

QY 200 -DEPSFQIRDYVESQKRSSCC 220  
Db 181 GERPNLKI-DSTPVKASGGCC 201

RESULT 13  
US-09-794-257-16  
Sequence 16, Application US/09794257  
Patent No. US20020009804A1  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel  
TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1  
FILE OF INVENTION: Human G-Proteins  
FILE REFERENCE: 35800/209285  
CURRENT APPLICATION NUMBER: US/09/794,257  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: 60/185,606  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Pfam accession number PF00071  
US-09-794-257-16

Query Match 40.5%; Score 466; DB 10; Length 198;  
Best Local Similarity 47.1%; Pred. No. 7.4e-43;  
Matches 97; Conservative 38; Mismatches 45; Indels 26; Gaps 6;

QY 31 KVMLLDGTGVTCTFLIQFDGAFSLGCTFIATVGDIFRNKVVTVVDGVRVKIQIWDTAGOE 90  
Db 1 KLVLLGDSGVGKSCLLLRFDKNFVE-EYPTIGVDFTYKTVVDGKTVKLVQIWDTAGOE 59

QY 91 RFRSVTHAYYRDAQALLLLYDITNKSFFDNIRAWLTIHEHYAQRDVVIMLLGNKADMS 148  
Db 60 RFRALPAYRCAQGFLLVYDITSRDSFENKVKWLEILRHADKDNVPIVLGNKCDLE 119

QY 149 SE-----RVIRSEGETLAREYG-VPEFLETSAKTGMNVELAFALAIKELKYRAGH 197  
Db 120 DDELDLELREGQKRVVSTEGEALAKELGALPEMETSAKTNTNVEEAFEELKVKVSE 179

QY 198 ---QADEPSFQIRDYVESOKKRSSCC 220  
Db 180 VVNVLDQP-----AKKKSKCC 196

## RESULT 14

US-09-945-173-5  
; Sequence 5, Application US/09945173  
; Patent No. US20020127568A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 38155-20035.00  
; CURRENT APPLICATION NUMBER: US/09/945,173  
; CURRENT FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/229,293  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Consensus amino acid sequence  
US-09-945-173-5

Query Match 40.5%; Score 466; DB 10; Length 198;  
Best Local Similarity 47.1%; Pred. No. 7.4e-43;  
Matches 97; Conservative 38; Mismatches 45; Indels 26; Gaps 6;

QY 31 KVMLLGDTGVGKTCFLIQFDGAFSLSGTFTATVGIDFRNKKVVTVDGVRVKLQIWDTAGQE 90  
Db 1 KLVLLGDSGVGKSSLLIRFTDNKFE-EYIPTIGVDFTYKTVEVDGKTVKLVQIWDTAGQE 59  
QY 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLFEIHEYAQRD--VVIMLLGNKADMS 148  
Db 60 RFRALRPAYYRGAQGFLLVDITSRDSFENVKKWLEELRHADKDNVPIVLVGNKCDLE 119  
QY 149 SE-----RVIRSEDEGETLAREYG-VPFLETSAKTGMNVVELAFLAIKELKYRAGH 197  
Db 120 DDEDLTEGQKRVVSTEGEALAKELGALPFMTSAKTNTNVEEAFEELAREILKKVSE 179  
QY 198 ---QADEPSFQIRDYVESOKKRSSCC 220  
Db 180 VVNVLDQP-----AKKKSKCC 196

## RESULT 15

US-09-972-529-4  
; Sequence 4, Application US/09972529  
; Patent No. US20020150916A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 47316, A NOVEL HUMAN G-PROTEIN AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 38155-20041.00  
; CURRENT APPLICATION NUMBER: US/09/972,529  
; CURRENT FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US 60/237,716  
; PRIOR FILING DATE: 2000-10-05  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Consensus amino acid sequence

US-09-972-529-4

Query Match 40.5%; Score 466; DB 10; Length 198;  
Best Local Similarity 47.1%; Pred. No. 7.4e-43;  
Matches 97; Conservative 38; Mismatches 45; Indels 26; Gaps 6;

QY 31 KVMLLGDTGVGKTCFLIQFDGAFSLSGTFTATVGIDFRNKKVVTVDGVRVKLQIWDTAGQE 90  
Db 1 KLVLLGDSGVGKSSLLIRFTDNKFE-EYIPTIGVDFTYKTVEVDGKTVKLVQIWDTAGQE 59  
QY 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLFEIHEYAQRD--VVIMLLGNKADMS 148  
Db 60 RFRALRPAYYRGAQGFLLVDITSRDSFENVKKWLEELRHADKDNVPIVLVGNKCDLE 119  
QY 149 SE-----RVIRSEDEGETLAREYG-VPFLETSAKTGMNVVELAFLAIKELKYRAGH 197  
Db 120 DDEDLTEGQKRVVSTEGEALAKELGALPFMTSAKTNTNVEEAFEELAREILKKVSE 179  
QY 198 ---QADEPSFQIRDYVESOKKRSSCC 220  
Db 180 VVNVLDQP-----AKKKSKCC 196

Search completed: November 17, 2002, 16:43:36  
Job time : 21 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 13:58:08 ; Search time 2433 Seconds  
(without alignments)  
17799.725 Million cell updates/sec

Title: US-09-817-199A-1  
Perfect score: 2674  
Sequence: 1 ttccctgcggcgccgcaact.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: gb\_gss.\*
- 18: em\_gss\_hum.\*
- 19: em\_gss\_inv.\*
- 20: em\_gss\_pln.\*
- 21: em\_gss\_vrt.\*
- 22: em\_gss\_fun.\*
- 23: em\_gss\_mam.\*
- 24: em\_gss\_mus.\*
- 25: em\_gss\_othr.\*
- 26: em\_gss\_pro.\*
- 27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	921.6	34.5	1014	14	BM921365
2	884.2	33.1	887	9	AL522282
3	847.4	31.7	855	9	AL580781
4	802.6	30.0	833	9	AL522281
5	755.4	28.2	881	12	BG759655
6	741.6	27.7	868	14	BQ690583

7	733.2	27.4	761	13	BI819064
C	8	702	26.3	702	BM981939
	9	701.2	26.2	1051	BG283602
C	10	696.8	26.1	700	BQ028124
	11	643.6	24.1	1069	BG282782
C	12	626.4	23.4	629	BQ183276
	13	605.4	22.6	607	AF188522
C	14	601.6	22.5	740	BI767046
	15	600.6	22.5	691	BG253976
C	16	585.6	21.9	749	AL559085
	17	581.2	21.7	886	BQ686932
C	18	553.2	20.7	676	BB598938
	19	539.2	20.2	603	AW955318
C	20	529	19.8	529	BM706070
	21	527.6	19.7	651	BB633978
C	22	516.8	19.3	655	BB635649
	23	514.6	19.2	522	BM151643
C	24	503.4	18.8	507	BM151381
	25	473.6	17.7	480	AW270880
C	26	464.4	17.4	493	AA324998
	27	460.6	17.2	488	BM149118
C	28	442.4	16.5	676	BB206788
	29	440.4	16.5	766	BQ188402
C	30	423.8	15.8	553	BE755280
	31	422.4	15.8	733	BB619625
C	32	415.2	15.5	437	W36293
	33	412.4	15.4	793	BG256040
C	34	409.4	15.3	434	AA806096
	35	399.4	14.9	401	AW768344
C	36	397.4	14.9	657	BB196489
	37	385	14.4	403	AA631787
C	38	381	14.2	401	AA878511
	39	380.4	14.2	455	AW464449
C	40	377.8	14.1	385	AI243836
	41	374.2	14.0	380	BE241772
C	42	362.8	13.6	579	BB621927
	43	358	13.4	368	BM152052
C	44	347	13.0	348	AA310682
	45	345.6	12.9	404	BM149006

ALIGNMENTS

RESULT 1  
BM921365  
LOCUS AGENCOURT\_6626159 NIH\_MGC\_115 Homo sapiens cdna clone IMAGE:5752779  
DEFINITION 5', mRNA sequence.  
ACCESSION BM921365  
VERSION BM921365.1 GI:19371744  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1014)  
NIH-MGC <http://mgc.nci.nih.gov/>  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12787 row: i column: 04  
High quality sequence stop: 735.  
Location/Qualifiers 1..1014  
FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:575279"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector: pcMV-SPORT6; Site\_1: Not1; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: This is a NIH\_MGC Library."  
BASE COUNT 237 a 288 c 295 g 194 t  
ORIGIN

Query Match 34.5%; Score 921.6; DB 14; Length 1014;  
Best Local Similarity 96.7%; Pred. No. 5.2e-93;  
Matches 973; Conservative 0; Mismatches 29; Indels 4; Gaps 3;  
QY 2 TGCCTCGCGGGCGGACATGCTCCTCGTCCAGGACATGACGGGACGCCAGCGGC 61  
Db 8 TCGCCCTGCGGGCGGACATGCTCCTCGTCCAGGACATGACGGGACGCCAGCGGC 67  
QY 62 CATTGCCACCGGGATGGGAGGCGCCCGGAGGCTCCCGCCCTGCAGTCCGAGCTACGA 121  
Db 68 CGTTGCCACCGGGATGGGAGGCGCCCGGAGGCTCCCGCCCTGCAGTCCGAGCTACGA 127  
QY 122 CTTACGGGACAGGTGATGCTTCTGGGAGACACAGCGGTGGGAAACATGTTTCTGAT 181  
Db 128 CTTACGGGACAGGTGATGCTTCTGGGAGACACAGCGGTGGGAAACATGTTTCTGAT 187  
QY 182 CCAATTCAGAGCGGGCTCTCTGTCGGAACCTTCATAGCCACCGTGGCATAGATT 241  
Db 188 CCAATTCAGAGCGGGCTCTCTGTCGGAACCTTCATAGCCACCGTGGCATAGATT 247  
QY 242 CAGGAACAGGTGATGCTGATGGCTGAGAGTGAAGCTGCAGATCGGACACCGC 301  
Db 248 CAGGAACAGGTGATGCTGATGGCTGAGAGTGAAGCTGCAGATCGGACACCGC 307  
QY 302 TGGGACAGAGCGTTCCGAGCGTCAACCATGTTATACAGAGATGCTCAGCGCTTGCT 361  
Db 308 TGGGACAGAGCGTTCCGAGCGTCAACCATGTTATACAGAGATGCTCAGCGCTTGCT 367  
QY 362 TCTGCTGTATGACATCACCACAAATCTTCTTCGACACATCAGGCGCTGGCTCACTGA 421  
Db 368 TCTGCTGTATGACATCACCACAAATCTTCTTCGACACATCAGGCGCTGGCTCACTGA 427  
QY 422 GATTCATGATATGCCAGAGGACGTGGTATGATGCTGCTAGGCAACAAGCGGATAT 481  
Db 428 GATTCATGATATGCCAGAGGACGTGGTATGATGCTGCTAGGCAACAAGCGGATAT 487  
QY 482 GAGCAGCAAGAGTATCGTTCCGACAGGAGACCTTGGCCAGGAGTACGGTGT 541  
Db 488 GAGCAGCAAGAGTATCGTTCCGACAGGAGACCTTGGCCAGGAGTACGGTGT 547  
QY 542 TCCCTTCTCTGGACACGCGCAAGACTGGCATGATGTTGGAGTTAGCCTTTCTGGCCAT 601  
Db 548 TCCCTTCTCTGGACACGCGCCAGACTGGCATGATGTTGGAGTTAGCCTTTCTGGCCAT 607  
QY 602 CGCAAGGAACTGAATACCGGCGCGGATCAGGCGGATGAGCCAGCTTCCAGATCCG 661  
Db 608 CGCAAGGAACTGAATACCGGCGCGGATCAGGCGGATGAGCCAGCTTCCAGATCCG 667  
QY 662 AGACTATGATATCCAGAGAGCGCTCCAGCTGCTCTCTGATGTAATCCAGG 721  
Db 668 AGACTATGATATCCAGAGAGCGCTCCAGCTGCTCTCTGATGTAATCCAGG 727  
QY 722 GGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGACGCTTCCCGCTCCAGGCGCTGGCT 781  
Db 728 GGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGACGCTTCCCGCTCCAGGCGCTGGCT 787

QY 782 TATTCACAGAGGCTGAGCCAAT-GGGGAGAAAGATGGAGACTACTGTCACAGCCGCTTC 840  
Db 788 TATTCACAGAGGCTGAGCCAATGGGGAGAAAGATGGAGGACTACTGTCACAGCCGCTTC 847  
QY 841 CTAGCAGGAGGTATACCTCAACTCTGACTTGAGTTCTTCGGTCTCCCGCAT--CCAC 898  
Db 848 CTAGCAGGAGGTATACCTCAACTCTGACTTGAGTTCTTCGGGTTCTCCCGCATTCACCC 907  
QY 899 AGGGAGGGTAAACACTTACCTTATTTTATAGTACATAATTAATACCAAAAAA-GG 957  
Db 908 GGAAGGGTAAACACTTACCTTATTTTATAGTACATAATTAATACCAAAAAAAGG 967  
QY 958 CGCCTGGATCCCAAAAAACCGAGGCTGGAGCTAGTGGCCCTTTT 1003  
Db 968 CGCCTGGATCCCAAAAAACCGAGGCTGGAGCTAGTGGCCCTTT 1013

RESULT 2  
AL522282  
LOCUS  
DEFINITION AL522282 LTI\_NFL004\_NBC2 887 bp mRNA linear EST 13-FEB-2001  
prime, mRNA sequence.  
ACCESSION AL522282  
VERSION AL522282.1 GI:12785775  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 887)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source  
1..887  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DB008YF08"  
/clone\_lib="LTI\_NFL004\_NBC2"  
/sex="male"  
/tissue.type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a Noll-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 203 a 256 c 257 g 168 t 3 others  
ORIGIN

Query Match 33.1%; Score 884.2; DB 9; Length 887;  
Best Local Similarity 99.5%; Pred. No. 7.4e-89;  
Matches 883; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 15 GGCACCTGCTCCTCGTCCAGGACATGACGGGACGCGCGGCGGTGGCCACCCGG 74  
Db 1 GGCACCTGCTCCTCGTCCAGGACATGACGGGACGCGCGGCGGTGGCCACCCGG 60  
QY 75 GATGGCAGGCGCCCGGAGGCGCTCCCGCCCTGCAGTCCGAGCTACGACCTCAGGGCAAG 134  
Db 61 GATGGCAGGCGCCCGGAGGCGCTCCCGCCCTGCAGTCCGAGCTACGACCTCAGGGCAAG 120  
QY 135 GTGATCCTCTCTGGGAGACACAGCGCTCGGCAAAACATGTTCTCTGATCAATCAAGAC 194

Db	121	GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCCTGATCCAATTCAAAGAC	180
Qy	195	GGGGCCCTTCCTGTCGCGAACCCTTCATAGCCACCCCTCGCATAGACTTCAGGAACAAGGTG	254
Db	181	GGGGCCCTTCCTGTCGCGAACCCTTCATAGCCACCCCTCGCATAGACTTCAGGAACAAGGTG	240
Qy	255	GTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGAACGG	314
Db	241	GTGACTGTGGATGGCGTGAAAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGAACGG	300
Qy	315	TTCCGAAGCGTACCCATGCTTATTACAGAGATGCTCAGGCGCTTGCCTCTCTGTATGAC	374
Db	301	TTCCGAAGCGTACCCATGCTTATTACAGARATGCTCAGGCGCTTGCCTCTCTGTATGAC	360
Qy	375	ATCACCAACAATCTCTTCGCAACAATCAGGCGCTGGCTCACTGAGATTCATGAGTAT	434
Db	361	ATCACCAACAATCTCTTCGCAACAATCAGGCGCTGGCTCACTGAGATTCATGAGTAT	420
Qy	435	GCCAGAGGGACGTGGTGATCATGCTGTAGGCAACAAGCGGATATGAGCAGCGAAAGA	494
Db	421	GCCAGAGGGACGTGGTGATCATGCTGTAGGCAACAAGCGGATATGAGCAGCGAAARA	480
Qy	495	GTGATCGTTCCGAAGAGGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCCTGGAG	554
Db	481	GTGATCGTTCCGAAGAGGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCCTGGAG	540
Qy	555	ACCAGCGCCAGACTGCGCATGAATGTGGAGTTAGCCTTTCGSCCATCGCCCAAGAACTG	614
Db	541	ACCAGCGCCAGACTGCGCATGAATGTGGAGTTAGCCTTTCGSCCATCGCCCAAGAACTG	600
Qy	615	AAATACCGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG	674
Db	601	AAATACCGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG	660
Qy	675	TCCAGAGAAGCGCTCAGCTGCTGCTCTTCATGTGAATCCAGGGGGCAGAGAGAG	734
Db	661	TCCAGAGAAGCGCTCAGCTGCTGCTCTTCATGTGAATCCAGGGGGCAGAGAGAG	720
Qy	735	GCTCTGAGGACACAGGATCGACCTTCCCGCTCCAGCGCTGGCTTATTCAGAGGC	794
Db	721	GCTCTGAGGACACAGGATCGACCTTCCCGCTCCAGCGCTGGCTTATTCAGAGGC	780
Qy	795	TGAGCCAAATGGGGAAGATGGAGGACTCACTGCACAGCGCGCTTCCTAGCAGGAGCTA	854
Db	781	TGAGCCAAATGGGGAAGATGGAGGACTCACTGCACAGCGCGCTTCCTAGCAGGAGCTA	840
Qy	855	TACTCCAACTCTACTTGAGTTCTGCGGTCTCCCGCATCCACAGG	901
Db	841	TACTCCAACTCTACTTGAGTTCTGCGGTCTCCCGCATCCACAGG	887
RESULT 3			
AL580781/c			
LOCUS	AL580781	LTI_NFL008_Tc2	855 bp mRNA linear EST 16-FEB-2001
DEFINITION	AL580781	LTI_NFL008_Tc2	Homo sapiens cDNA clone CS0DJ014YB11 3
ACCESSION	AL580781		prime, mRNA sequence.
VERSION	AL580781		
KEYWORDS	AL580781.1	GI:12947137	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 855)		
COMMENT	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
FEATURES			

source		1. .855	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="CS0DJ014YB11"	
		/sex="male"	
		/tissue_type="T cells from T cell leukemia"	
		/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"	
BASE COUNT	204 a	224 c	224 g
ORIGIN			201 t
Query Match			
Best Local Similarity 31.7%; Score 847.4; DB 9; Length 855;			
Matches 848; Conservative 1; Mismatches 2; Indels 0; Gaps 0;			
Qy	1700	ATTTTTCGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTC	1759
Db	855	ATTTTTCGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTC	796
Qy	1760	AAGCAACCTGCGGCTCGGCCCTCCCAAAGTACTGGGATTTACACGACAGGACCATGTC	1819
Db	795	AAGCAACCTGCGGCTCGGCCCTCCCAAAGTACTGGGATTTACACGACAGGACCATGTC	736
Qy	1820	CCAGCTAGATGTGCTTATCCCAATCTTTGGCAGGATGCAGCTCCACAGGCGATTTC	1879
Db	735	CCAGCTAGATGTGCTTATCCCAATCTTTGGCAGGATGCAGCTCCACAGGCGATTTC	676
Qy	1880	TTCAGCAGCTGAAGTGTTCAGCCCTCTGGGTTAAGAGCCAGATGAAGAGAGAAATCCCTT	1939
Db	675	TTCAGCAGCTGAAGTGTTCAGCCCTCTGGGTTAAGAGCCAGATGAAGAGAGAAATCCCTT	616
Qy	1940	TCCTAGTTTGAATGTGTTGTAAGAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGG	1999
Db	615	TCCTAGTTTGAATGTGTTGTAAGAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTGG	556
Qy	2000	GAGACAAGATTAAAGCAAACTCCCTCGATGTATCTCTTTCGACCCCAAGCTCTGCCTCC	2059
Db	555	GAGACAAGATTAAAGCAAACTCCCTCGATGTATCTCTTTCGACCCCAAGCTCTGCCTCC	496
Qy	2060	TCCCTGACACCCATGCGCTTTCTTTAACTTCTCAAAACAGATACAGGCGCTAAACTGC	2119
Db	495	TCCCTGACACCCATGCGCTTTCTTTAACTTCTCAAAACAGATACAGGCGCTAAACTGC	436
Qy	2120	TTTACCTCCCTCTCTAGTGCAGTTAGTGGTGGAGGTCACCCATTTCCGAGTTAA	2179
Db	435	TTTACCTCCCTCTCTAGTGCAGTTAGTGGTGGAGGTCACCCATTTCCGAGTTAA	376
Qy	2180	ACCAATGCAATATGATGTAACAAAGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2239
Db	375	ACCAATGCAATATGATGTAACAAAGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG	316
Qy	2240	GCAAGTTCTATGTGCTCTCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGA	2299
Db	315	GCAAGTTCTATGTGCTCTCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGA	256
Qy	2300	AGTGACAGGAAACATGAGTGCATGACCTGCAGGCGATCTTTACTGCAGCTCTGCGGCT	2359
Db	255	AGTGACAGGAAACATGAGTGCATGACCTGCAGGCGATCTTTACTGCAGTCTCTGCGGCT	196
Qy	2360	GGAGGGGAGAGGGGAGGAGAGTATGCGCTGCACATTTCTGAGGCTTACTGCAATTTC	2419
Db	195	GGAGGGGAGAGGGGAGGAGAGTATGCGCTGCACATTTCTGAGGCTTACTGCAATTTC	136
Qy	2420	TTTCAAGGCAGAAATCTTGCTGTGAGCAGTCAGCGGCTCCAGTTTGGGCCCCGATAAGGA	2479

```
Db 135 TTTCAAGGCAGAAATCTTCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCCATAGGAA 76
|||||
Qy 2480 GTTCTCCGTGGCTCCTCAGGACAGCAGGAGGAGGCTGACATTGCCAGTCTCTTCTG 2539
|||||
Db 75 GTTCTCCGTGGCTCCTCAGGACAGCAGGAGGAGGCTGACATTGCCAGTCTCTTCTG 16
|||||
Qy 2540 GGGCCCAAGGC 2550
|||||
Db 15 GGGCCCAAGGC 5
|||||

RESULT 4
AL522281/c 833 bp mRNA linear EST 13-FEB-2001
LOCUS AL522281 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 3
DEFINITION prime, mRNA sequence.
ACCESSION AL522281
VERSION AL522281.1 GI:12785774
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 833)
Li.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. .833
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DB008YF08"
/tissue_type="neuroblastoma cells"
/lab_host="PH108"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: filiang@lifetech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 206 a 207 c 215 g 200 t 5 others
ORIGIN

Query Match 30.0%; Score 802.6; DB 9; Length 833;
Best Local Similarity 98.7%; Pred. No. 7.3e-80;
Matches 814; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 1673 GTACCATCACCTGGCTAATTTTGTATTTTGTAGACACAGGGTTTCGCCATCTTGC 1732
|||||
Db 825 KTACCATCACCTGGCTAATTTTGTATTTTGTAGACACAGGGTTTCGCCATCTTGC 766
|||||
Qy 1733 CCAGGCTGGTCTTGAATCTCTGAGCTCAAGCAACCTGCCGGCTCGGCCTCCCAAGTAC 1792
|||||
Db 765 CCAGGCTGGTCTTGAATCTCTGAGCTCAAGCAACCTGCCGGCTCGGCCTCCCAAGTAC 706
|||||
Qy 1793 TGGATTACACGACGAGGACCATGCCAGGCTAGATGTCTTATCCCAATCCTTTGG 1852
|||||
Db 705 TGGATTACACGACGAGGACCATGCCAGGCTAGATGTCTTATCCCAATCCTTTGG 646
|||||
Qy 1853 CAGGCATGACGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTATAGCCCTCCCTGGGT 1912
|||||
Db 645 CAGGCATGACGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTATAGCCCTCCCTGGGT 586
|||||
```

```
Qy 1913 TAAGAGCCAGATAAGAGAGAAATCCCTTTCTTAGGTTTGAATGTCTTGTGAAAAAAGA 1972
|||||
Db 585 TAAGAGCCAGATAAGAGAGAAATCCCTTTCTTAGGTTTGAATGTCTTGTGAAAAAAGA 526
|||||
Qy 1973 GAAATCCCTGGCTCCTGGAGCTGGTGGAGACAAGATTAAAGCAAACTCCCTCGACATGT 2032
|||||
Db 525 GAAATCCCTGGCTCCTGGAGCTGGTGGAGACAAGATTAAAGCAAACTCCCTCGACATGT 466
|||||
Qy 2033 ATCCCTTTGACCCCAAGCTCTGCTCTCCCTCCAGCACCAATGCCCTTTCCCTTTAACTTC 2092
|||||
Db 465 ATCCCTTTGACCCCAAGCTCTGCTCTCCCTCCAGCACCAATGCCCTTTCCCTTTAACTTC 406
|||||
Qy 2093 TCAACAGATACACAGGCGCTAAACTGCTTTTACCTCCCTCCCTACTG-AGTCAGGTTAGT 2151
|||||
Db 405 TCAACAGATACACAGGCGCTAAACTGCTTTTACCTCCCTCCCAAACTGCACTCAGGTTAGT 346
|||||
Qy 2152 GGTGGAGGTCAACCATTTCCGAGTTAAACCAATCAATATAGTAAACAAGTCATGT 2211
|||||
Db 345 GGTGGAGGTCAACCATTTCCGAGTTAAACCAATCAATATAGTAAACAAGTCATGT 286
|||||
Qy 2212 GGGTATGCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGCTCTCCCTTGGTCACATATC 2271
|||||
Db 285 GGGTATGCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGCTCTCCCTTGGTCACATATC 226
|||||
Qy 2272 TCCAAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGTCTGACCTGCA 2331
|||||
Db 225 TCCAAAAGCTCTGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGTCTGACCTGCA 166
|||||
Qy 2332 GGCATCTTTACTGCAGCTCTGCCGCTGGAGGGGAGAGGAGGAAGAGTATGCC 2391
|||||
Db 165 GGCATCTTTACTGCAGCTCTGCCGCTGGAGGGGAGAGGAGGAAGAGTATGCC 106
|||||
Qy 2392 TGCACATTTCTGAGGCTACTGCATTTGCTTCAAGSCAGAAATCTTGCTCTGAGCAGTCA 2451
|||||
Db 105 TGCACATTTCTGAGGCTACTGCATTTGCTTCAAGSCAGAAATCTTGCTCTGAGCAGTCA 46
|||||
Qy 2452 GGGGCTCCAGTTTGGCCCGGATAAGGAAGTTCTCGTGGCCCTCC 2496
|||||
Db 45 GCGGCTCCAGTTTGGCCCGGATAAGGAAGTTCTCGGTCGCCCTCCV 1
|||||

RESULT 5
Bg759655 881 bp mRNA linear EST 15-MAY-2001
LOCUS 602713348F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853626 5',
DEFINITION mRNA sequence.
ACCESSION Bg759655
VERSION Bg759655.1 GI:14070308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1699 row: p column: 11
High quality sequence stop: 707.
Location/Qualifiers
1. .881
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
```



```
/clone="IMAGE:4853626"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/site="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      220 a      236 c      261 g      164 t
ORIGIN

Query Match      28.2%; Score 755.4; DB 12; Length 881;
Best Local Similarity 94.6%; Pred. No. 1.1e-74;
Matches 805; Conservative 0; Mismatches 41; Indels 5; Gaps 2;

Qy 566 GACTGGCATGAATGTGAGTTAGCTTTCTGGCCATCGCCAAAGCACTGAAATACCGGC 625
Db      |||||||
Qy 2 GACTGGCATGAATGTGAGTTAGCTTTCTGGCCATCGCCAAAGCACTGAAATACCGGC 61
Db      |||||||

Qy 626 CGGSCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAA 685
Db      |||||||

Qy 62 CGGSCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAA 121
Db      |||||||

Qy 686 GCGTTCAGCTGCTGCTTCATGTGATCCAGGCGGAGAGAGGCTCTGGAGGC 745
Db      |||||||

Qy 122 GCGTTCAGCTGCTGCTTCATGTGATCCAGGCGGAGAGAGGCTCTGGAGGC 181
Db      |||||||

Qy 746 ACACAGATGACAGCTTCCCTCCAGGCTGCTTATCCAGAGGCTGAGCAATGG 805
Db      |||||||

Qy 182 ACACAGATGACAGCTTCCCTCCAGGCTGCTTATCCAGAGGCTGAGCAATGG 241
Db      |||||||

Qy 806 GGAGAAAGATGGAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATACTCAACTC 865
Db      |||||||

Qy 242 GGAGAAAGATGGAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATACTCAACTC 301
Db      |||||||

Qy 866 CTACTTGAATGCTGCGCTTCCCGCATCCACAGGAGGTTAAACACTTAGTTTAT 925
Db      |||||||

Qy 302 CTACTTGAATGCTGCGCTTCCCGCATCCACAGGAGGTTAAACACTTAGTTTAT 361
Db      |||||||

Qy 926 TTTAATAGTACATAATTTAATACAAAAGCGCTTGATGCCCAAAAACCCAGGCTG 985
Db      |||||||

Qy 362 TTTAATAGTACATAATTTAATACAAAAGCGCTTGATGCCCAAAAACCCAGGCTG 421
Db      |||||||

Qy 986 GGAGCTAGTGGCCCTTTTGTCTTAGGACTTGGGGGCGCGCCCTCCTCTAAGCATA 1045
Db      |||||||

Qy 422 GGAGCTAGTGGCCCTTTTGTCTTAGGACTTGGGGGCGCGCCCTCCTCTAAGCATA 481
Db      |||||||

Qy 1046 ACAAGGTGGTGTGCTCCAGCTCAGCCCGACAGGGGACACAGATGCATTTGGGGGTGAGG 1105
Db      |||||||

Qy 482 ACAAGGTGGTGTGCTCCAGCTCAGCCCGACAGGGGACACAGATGCATTTGGGGGTGAGG 541
Db      |||||||

Qy 1106 GCAGGTAATGACTCATCGACCTCAGTTCAGCTGGACAGGCTCAGGTGACCCAGC 1165
Db      |||||||

Qy 542 GCAGGTAATGACTCATCGACCTCAGTTCAGCTGGACAGGCTCAGGTGACCCAGC 601
Db      |||||||

Qy 1166 CTTTCACTGTCTCCGCTCTCCAGAGCTTATTTTCGGCCCATCTCCCAAAATAGTGGGCC 1225
Db      |||||||

Qy 602 CTTTCACTGTCTCCGCTCTCCAGAGCTTATTTTCGGCCCATCTCCCAAAATAGTGGGCC 661
Db      |||||||

Qy 1226 CTTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATGAAGATATGGAGATGGAGGGG 1285
Db      |||||||

Qy 662 CTTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATGAAGATATGGAGATGGAGGGG 721
Db      |||||||

Qy 1286 GAGCACAAAGGGGACAGAGTAGGCTCTA ---GCTGGCTATCTCTGGCCCTTACTTAACACC 1342
Db      |||||||

Qy 722 GAGGAAAGGGGACAGAGTAGGCTCTAAGGCAAGCATGGCAAGTCTCTGGCCGTACAAAGAGCG 781
Db      |||||||

Qy 1343 CCCTGG --AGGSCATGCCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTTGGAAATAT 1400
Db      |||||||
```

```
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 782 CCCTGGGAGGCGATGCCCTTATAAGCCAGGACAAGCCCATTTGGGCACCTTGGAAATAG 841
Qy 1401 TGGTTCAGGC 1411
Db 842 TGGACCGGAG 852

RESULT 6
BQ690583 BQ690583 868 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8186311 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251642
DEFINITION 5', mRNA sequence.
ACCESSION BQ690583
VERSION BQ690583.1 GI:21815899
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcgaps@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2396 row: k column: 03
High quality sequence stop: 532.
FEATURES
source
1..868
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6251642"
/clone_lib="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 215 a 239 c 240 g 172 t 2 others
ORIGIN

Query Match 27.7%; Score 741.6; DB 14; Length 868;
Best Local Similarity 93.6%; Pred. No. 3.5e-73;
Matches 805; Conservative 0; Mismatches 51; Indels 4; Gaps 3;

Qy 490 AAAGAGTGTATCGTTCGGAAGCAGGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCC 549
Db 1 AAAGAGTGTATCGTTCGGAAGCAGGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCC 60
Qy 550 TGGAGACCCAGCCCAAGACTGCATGAATGTGGAGTTAGCCCTTCTTGGCCATCGCAAGG 609
Db 61 TGGAGACCCAGCCCAAGACTGCATGAATGTGGAGTTAGCCCTTCTTGGCCATCGCAAGG 120
Qy 610 AACTGAAATACCGGCGCGGATCAGCGGATGAGCCAGCTTCCAGACTCCAGAGACTATG 669
Db 121 AACTGAAATACCGGCGCGGATCAGCGGATGAGCCAGCTTCCAGACTCCAGAGACTATG 180
Qy 670 TAGAGTCCCAAGAGAGCGCTCCAGCTGCTCTCTCATGTGAATCCAGGGGCGCAGAG 729
Db 181 TAGAGTCCCAAGAGAGCGCTCCAGCTGCTCTCTCATGTGAATCCAGGGGCGCAGAG 240
```

QY 730 AGGAGGCTCTGGAGGCACACAGGATGACAGCTTCCCTCCAGGCTGGCTTATTCCAA 789  
|||||  
Db 241 AGGAGGCTCTGGAGGCACACAGGATGACAGCTTCCCTCCAGGCTGGCTTATTCCAA 300  
QY 790 GAGGCTGAGCAATGGGGAAGAAGATGGAGGACTCACTGCACAGCGCTTCTACAGGG 849  
|||||  
Db 301 GAGGCTGAGCAATGGGGAAGAAGATGGAGGACTCACTGCACAGCGCTTCTACAGGG 360  
QY 850 AGCTATACCTCAACCTCTACTTGAGTCTCTGGGTCCTCCGCCATCCACAGGAGGTAA 909  
|||||  
Db 361 AGCTATACCTCAACCTCTACTTGAGTCTCTGGGTCCTCCGCCATCCACAGGAGGTAA 420  
QY 910 AACACTTAGCTTTTATTTAATAGTACATAATTTAATCAAAAAAGGCGCTGGATCCC 969  
|||||  
Db 421 AACACTTAGCTTTTATTTAATAGTACATAATTTAATCAAAAAAGGCGCTGGATCCC 480  
QY 970 CAAAAACCGAGCTGGGAGCTAGTGGCCCTTTTGGTTTCTAGGACTT-GGGGGGCGGC 1028  
|||||  
Db 481 CAAAAACCGAGCTGGGAGCTAGTGGCCCTTTTGGTTTCTAGGACTTGGGGGGCGGC 540  
QY 1029 CTTCCCTCTTAAGCATAAACAAAGGTGGTGTCTCCAGCTCAGCCCGGACACAGAT 1088  
|||||  
Db 541 CTTCCCTCTTAAGCATACNAAGTGGTGTCTCCAGCTCAGCCCGGACACAGAT 600  
QY 1089 GCACCTTT-GGGGGTGAGGCGAGGTAAATGACTCCATCGCACCTCAGTTCAGTGGACAGA 1147  
|||||  
Db 601 GCACCTTTGGGGGTGAGGCGAGGTAAATGACTCCATCGCACCTCAGTTCAGTGGACAGA 660  
QY 1148 GGTGAGGTGACCCAGCCTTCACTGTCTCCGCTCTCCAGGACTTATCTCGCCCAT 1207  
|||||  
Db 661 GGTGAGGTGACCCAGCCTTCACTGTCTCCGCTCTCCAGGACTTATCTTCGCCCAT 720  
QY 1208 CTCCTCAATTAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAG 1267  
|||||  
Db 721 CTCCTCAATTAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAATTAAGA 780  
QY 1268 ATATGAGATGAGGAGGAGGACAAAGGG--CAGAGTAGGAGTGTAGCTGGCTATCTC 1325  
|||||  
Db 781 TATGAAAGTGAAGGGGGAAGCAAGGGGGCAAAAAGAGGCTTAACCTGGCTTATC 840  
QY 1326 TGGCCTTACTACACCCCTC 1345  
|||||  
Db 841 TCTTGGCCTTAACACACCC 860  
  
RESULT 7  
BI819064  
LOCUS 761 bp mRNA linear EST 04-OCT-2001  
DEFINITION 60303417f1 NIH\_MGC\_115 Homo sapiens cdna clone IMAGE:5174308 5',  
mRNA sequence.  
ACCESSION BI819064  
VERSION BI819064.1 GI:15930614  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 761)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cdna Library Preparation: Life Technologies, Inc.  
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1434 row: b column: 05  
High quality sequence stop: 759.

FEATURES  
source  
1. 761  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5174308"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector: pcwV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."  
BASE COUNT 174 a 219 c 204 g 164 t  
ORIGIN  
Query Match 27.4%; Score 733.2; DB 13; Length 761;  
Best Local Similarity 99.3%; Pred. No. 3.3e-72;  
Matches 757; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
QY 728 AGAGGAGGCTCTGGAGGCACACAGGATGAGCCTTCCCTCCAGGCTGGCTTATTCC 787  
|||||  
Db 1 AGAGGAGGCTCTGGAGGCACACAGGATGAGCCTTCCCTCCAGGCTGGCTTATTCC 60  
QY 788 AAGAGGCTGAGCAATG-GGGAGAAAGATGGAGGACTCACTGCACAGCGCTTCCTAGCA 846  
|||||  
Db 61 AAGAGGCTGAGCAATGTTGGGAGAAAGATGGAGGACTCACTGCACAGCGCTTCCTAGCA 120  
QY 847 GGGAGCTATACCTCAACTCTCTACTTCTCTGGGCTCTCCCGCATCCACAGGAGGG 906  
|||||  
Db 121 GGGAGCTATACCTCAACTCTCTACTTCTCTGGGCTCTCCCGCATCCACAGGAGGG 180  
QY 907 TAAACACTTAGCTTTTATTTAATAGTACATAATTTAATACCAAAAAGGCGCTGGAT 966  
|||||  
Db 181 TAAACACTTAGCTTTTATTTAATAGTACATAATTTAATACCAAAAAGGCGCTGGAT 240  
QY 967 CCCCCAAAAACCGAGCTGGGAGCTAGTGGCCCTTTTGTCTTCTAGGACTTGGGGGGCG 1026  
|||||  
Db 241 CCCCCAAAAACCGAGCTGGGAGCTAGTGGCCCTTTTGTCTTCTAGGACTTGGGGGGCCA 300  
QY 1027 GCCCTCCCTCTTAAGCATAAACAAGTGGTGTGTCTCCAGCTCAGCCCGGAGACACAG 1086  
|||||  
Db 301 GCCCTCCCTCTTAAGCATAAACAAGTGGTGTGTCTCCAGCTCAGCCCGGAGACACAG 360  
QY 1087 ATGCACCTTTGGGGTGGAGGAGTAATGACTCCATCGCACCTCAGTTCAGCTGGACAG 1146  
|||||  
Db 361 ATGCACCTTTGGGGTGGAGGAGTAATGACTCCATCGCACCTCAGTTCAGCTGGACAG 420  
QY 1147 AGGCTCAGGTGACCCAGCCTTCTACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCA 1206  
|||||  
Db 421 AGGCTCAGGTGACCCAGCCTTCTACTGTCTCCCGCTCTCCAGGAGCTTATCTTCGCCCA 480  
QY 1207 TCTCCAAAATAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGA 1266  
|||||  
Db 481 TCTCCAAAATAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGA 540  
QY 1267 GATATGGAGATGGAGGGGAGGACAAAGGGCAGAGAGTAGGCTTAGCTGGCTATCTCT 1326  
|||||  
Db 541 GATATGGAGATGGAGGGGAGGACAAAGGGCAGAGAGTAGGCTTAGCTGGCTATCTCT 600  
QY 1327 GGCCTTACTAACACCCCTCGGAGGATCCCTTTTCTCCAGCACACAAAGACATTTGG 1386  
|||||  
Db 601 GGCCTTACTAACACCCCTCGGAGGATCCCTTTTCTCCAGCACACAAAGACATTTGG 660  
QY 1387 GCACCTTGAATAATTTGGTTCCAGGCTTCTGTCTCTGAGTTCAGATCTCGGGGAGCCC 1446  
|||||  
Db 661 GCACCTTGAATAATTTGGTTCCAGGCTTCTGTCTCTGAGTTCAGATCTCGGGGAG-CC 719  
QY 1447 CTCCTCCCTCCCTGNAATCCCTGGGCTTAGCTACCTTCTCTGCTGT 1488

Db	720	CTCCCCCCTGAATCCCTGGCTTAGCTACCTTCCCTGCTGT 761		
RESULT 8				
BM981939/c				
LOCUS				
DEFINITION		702 bp mRNA linear EST 21-MAR-2002		
ACCESSION		UI-CF-EN1-adg-h-11-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone		
VERSION		BM981939		
KEYWORDS		UI-CF-EN1-adg-h-11-0-UI 3', mRNA sequence.		
SOURCE		BM981939.1 GI:19604935		
ORGANISM		human.		
REFERENCE		Homo sapiens		
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE		Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL		1 (bases 1 to 702)		
MEDLINE		Bonaldo, M.F., Lennon, G. and Soares, M.B.		
COMMENT		Normalization and subtraction: two approaches to facilitate gene discovery		
		Genome Res. 6 (9), 791-806 (1996)		
		9704477		
		Contact: McCray, PB		
		McCray Lab		
		University of Iowa		
		2024 University of Iowa Med Labs, Iowa City, IA 52242, USA		
		Tel: 319 356 4866		
		Fax: 319 356 7171		
		Email: paul-mccray@uiowa.edu		
		Tissue Procurement: Dr. M. J. Welsh, University of Iowa		
		cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
		CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
		DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
		Clone Distribution: Researchers may obtain clones from Research		
		Genetics (www.resgen.com).		
		Seq primer: M13 FORWARD		
		POLYA=Yes.		
FEATURES		Location/Qualifiers		
source		1..702		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="UI-CF-EN1-adg-h-11-0-UI"		
		/clone_lib="UI-CF-EN1"		
		/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"		
		/dev_stage="Adult"		
		/lab_host="DH10B (Life Technologies) (T1 phage resistant)"		
		/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.		
		TAG_LIB=UI-CF-EN1		
		TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS		
		6hr to LPS 24h		
		TAG_SEQ=CTGCTCAGGT		
BASE COUNT	166 a	181 c 171 g 184 t		
ORIGIN				
Query Match		26.3%; Score 702; DB 14; Length 702;		
Best Local Similarity		100.0%; Pred. No. 9.5e-69;		
Matches 702; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;		
QY	1953	ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAA 2012		

Plate: LLAM10417 row: k column: 08  
High quality sequence stop: 782.

FEATURES  
source

Location/Qualifiers  
1. .1051  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4520191"  
/clone\_lib="NIH\_MGC\_91"  
/tissue\_type="adrenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: prostate; Vector: pCMV-SPORT6; Site: 1. Not I;  
Site: 2. SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.4 Kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: This is a NIH\_MGC Library."

BASE COUNT 275 a 287 c 314 g 175 t  
ORIGIN

Query Match 26.2%; Score 701.2; DB 12; Length 1051;  
Best Local Similarity 99.3%; Pred. No. 8.3e-69;  
Matches 725; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 40 ACATGAGCGGCGCAGGCGCGTTCGCCACCGGATGCGAGGCCCGCGAGCGCTCCC 99  
DB 1 ACATGAGCGGCGCAGGCGCGTTCGCCACCGGATGCGAGGCCCGCGAGCGCTCCC 60  
QY 100 CGCCCTGCAGTCGAGCTACGACCTCACGGCAAGGTGATGCTTCCTGGGAGACACAGGCG 159  
DB 61 CGCCCTGCAGTCGAGCTACGACCTCACGGCAAGGTGATGCTTCCTGGGAGACACAGGCG 120  
QY 160 TCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCTTCCTGTCGGGAACCTTCA 219  
DB 121 TCGGCAAAACATGTTTCCTGATCCAATTCAAAGACGGGGCTTCCTGTCGGGAACCTTCA 180  
QY 220 TAGCCACGCTCGCATAGACTTCAGAAACAGGTGCTGACTGTGGATGGCTGAGAGTGA 279  
DB 181 TAGCCACGCTCGCATAGACTTCAGAAACAGGTGCTGACTGTGGATGGCTGAGAGTGA 240  
QY 280 AGCTGCAGATCTGGGACACCGCTGGGAGGAGCGTTCCGAAGCGTCACCATGCTTATT 339  
DB 241 AGCTGCAGATCTGGGACACCGCTGGGAGGAGCGTTCCGAAGCGTCACCATGCTTATT 300  
QY 340 ACAGAGATGCTCAGGCTTCCTGCTGATGACATCACCACCAATCTCTTTTCGACA 399  
DB 301 ACAGAGATGCTCAGGCTTCCTGCTGATGACATCACCACCAATCTCTTTTCGACA 360  
QY 400 ACATCAGGCGCTGGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATG 459  
DB 361 ACATCAGGCGCTGGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATG 420  
QY 460 TGCTAGCAACAAGCGGATATGAGCAGGAGAGAGTGTGCTTCGAGACGGAGAGA 519  
DB 421 TGCTAGCAACAAGCGGATATGAGCAGGAGAGAGTGTGCTTCGAGACGGAGAGA 480  
QY 520 CCTTGGCCAGGAGTACGGTGTTCCTTCTGAGACCGCGCCAGACATGGCATGAATG 579  
DB 481 CCTTGGCCAGGAGTACGGTGTTCCTTCTGAGACCGCGCCAGACATGGCATGAATG 540  
QY 580 TGGAGTTAGCTTCTTGGCCATGCCAAGAACTGAAATACCGGCGCGGCATCAGCGCG 639  
DB 541 TGGAGTTAGCTTCTTGGCCATGCCAAGAACTGAAATACCGGCGCGGCATCAGCGCG 600  
QY 640 ATGAGCCAGCTTCCAGATCCGAGATATGATAGATGCCAG-AAGAGCGCTCCAGCTGC 698  
DB 601 ATGAGCCAGCTTCCAGATCCGAGATATGATAGATGCCAGAAAGAGCGCTCCAGCTGC 660  
QY 699 TGCTCTTCATGTGAATCCAGGGG-CAGAGAGAGGCTCTCGGAGCACACAGGATGCA 757  
DB 661 TGCTCTTCATGTGAATCCAGGGGCGCAGAGAGAGGCTCTCGGAGCACACAGGATGCA 720  
QY 758 GCCTTCCCC 767  
DB 721 GCCTTCCCC 730

RESULT 10  
B0028124/cLOCUS  
DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

B0028124 700 bp mRNA linear EST 27-MAR-2002  
UI-H-C00-ari-d-07-0-UI.s1 NCI\_CGAP\_Sub9 Homo sapiens cDNA clone  
IMAGE:3106595 3', mRNA sequence.

B0028124  
B0028124.1 GI:19763403  
EST.  
human.

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 700)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)

Tissue procurement: Dr. Jose Mercuende  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Seq primer: M13 FORWARD  
POLYA=Yes.

Location/Qualifiers  
1. .700  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3106595"  
/clone\_lib="NCI\_CGAP\_Sub9"  
/tissue\_type="mixed"  
/vet\_stage="mixed"

/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73-Pac (Pharmacia) with a modified  
polylinker; Site: 1. EcoR I; Site 2: Not I; tissues:  
Cholonic mucosa with Crohns disease, Cholonic mucosa with  
ulcerative colitis, Fetal thymus, Cervix, Cervical  
adenosquamous carcinoma, Ligament cells, Prostate  
carcinoma, Bladder carcinoma, Brain oligodendrocyte ;  
NCI\_CGAP\_Sub9 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are CGTC, AACG, GGGCC,  
GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional  
information, contact: Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)

BASE COUNT 166 a 182 c 170 g 182 t  
ORIGIN

Query Match 26.1%; Score 696.8; DB 14; Length 700;  
Best Local Similarity 99.7%; Pred. No. 3.6e-68;  
Matches 698; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1950 GGAATGTGTGAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGAGACAGAT 2009  
DB 700 GGAATGTGTGAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGAGACAGAT 641

QY 2010 TAAGCAAACTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCA 2069  
DB 640 TAAGCAAACTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCA 581

```
QY 2070 CCCATGCCCTTCCCTTAACTTCTCAACAGATACACAGGCCCTAAACTGCTTTACTCCTCC 2129
|||||
Db 580 CCCATGCCCTTCCCTTAACTTCTCAACAGATACACAGGCCCTAAACTGCTTTACTCCTCC 521
QY 2130 CTCCTACTGAGTCAGGTTAGGTGGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAA 2189
|||||
Db 520 CTCCTACTGAGTCAGGTTAGGTGGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAA 461
QY 2190 TATGAGTAAACAAAGTCATGTGGGTATCTCTGGGTAGAGAGGGGTAGCAAGTTTCA 2249
|||||
Db 460 TATGAGTAAACAAAGTCATGTGGGTATCTCTGGGTAGAGAGGGGTAGCAAGTTTCA 401
QY 2250 GTGTCTCTCTGGTCACATATCTCCAAAGCTCTGATCCCTGCGCATGGGAAGTGGACAGG 2309
|||||
Db 400 GTGTCTCTCTGGTCACATATCTCCAAAGCTCTGATCCCTGCGCATGGGAAGTGGACAGG 341
QY 2310 AAACATGAGGTCATGACCTGACGAGGATCTTTACTGAGCTCTGCGGCGCTGAGAGGGGAG 2369
|||||
Db 340 AAACATGAGGTCATGACCTGACGAGGATCTTTACTGAGCTCTGCGGCGCTGAGAGGGGAG 281
QY 2370 AGGGGAGGAAGATGATCGCTGCACATTTCTGAGGCTACTGAGGCTTCTGCTTTCAAGGCA 2429
|||||
Db 280 AGGGGAGGAAGATGATCGCTGCACATTTCTGAGGCTACTGAGGCTTCTGCTTTCAAGGCA 221
QY 2430 GAAATCTTCTCTGAGCAGCTCAGCGGCTCCAGTTTGGGCGCGATAGGAAGTTCTCCGCTG 2489
|||||
Db 220 GAAATCTTCTCTGAGCAGCTCAGCGGCTCCAGTTTGGGCGCGATAGGAAGTTCTCCGCTG 161
QY 2490 GCGTCCCTCAGGAGGAGGAGGAGGCTGACATTCGACGTCCTTCCTGGGGCCCAAGG 2549
|||||
Db 160 GCGTCCCTCAGGAGGAGGAGGAGGCTGACATTCGACGTCCTTCCTGGGGCCCAAGG 101
QY 2550 CAGGTTCAGGAGATCCATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTCA 2609
|||||
Db 100 CAGGTTCAGGAGATCCATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTCA 41
QY 2610 GAATTAACCTGAGTATTTTGGAAAGCAAAAAA 2649
|||||
Db 40 GAATTAACCTGAGTATTTTGGAAAGCAAAAAA 1
```

```
RESULT 11
LOCUS BG282782 1069 bp mRNA linear EST 21-FEB-2001
DEFINITION 502405867F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4518136 5',
mRNA sequence.
```

```
ACCESSION BG282782
VERSION BG282782.1 GI:13031999
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1069)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DPD
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10412 row: e column: 17
High quality sequence stop: 689.
```

```
FEATURES
source
1..1069
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

```
/clone="IMAGE:4518136"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell_line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 264 a 296 c 295 g 214 t
ORIGIN
```

```
Query Match 24.1%; Score 643.6; DB 12; Length 1069;
Best Local Similarity 94.3%; Pred. No. 1.7e-62;
Matches 722; Conservative 0; Mismatches 39; Indels 5; Gaps 5;
QY 1104 GGGCAGGTAATGACTCCATCGCACCCCTCAGTTTCAGGTGACAGAGGCTCAGGTGACCCCA 1163
|||||
Db 1 GGGCAGGTAATGACTCCATCGCACCCCTCAGTTTCAGGTGACAGAGGCTCAGGTGACCCCA 60
QY 1164 GCCTTCACCTGCTCCCGCTCTCCAGGAGCTTATCTTCGCCCATCTCCCAATAAGTGGG 1223
|||||
Db 61 GCCTTCACCTGCTCCCGCTCTCCAGGAGCTTATCTTCGCCCATCTCCCAATAAGTGGG 120
QY 1224 CCCTTGTGCTGTGAGGAGACCAAGCCCTCAGGGAAGATAAGAGATATGAGATGGGAGG 1283
|||||
Db 121 CCCTTGTGCTGTGAGGAGACCAAGCCCTCAGGGAAGATAAGAGATATGAGATGGGAGG 180
QY 1284 GGGAGGACAAAGGGGACAGAGTAGGTCTAGCTGGCTATCTCTGGCCTTACTTAACACCC 1343
|||||
Db 181 GGGAGGACAAAGGGGACAGAGTAGGTCTAGCTGGCTATCTCTGGCCTTACTTAACACCC 240
QY 1344 CCTGAGGCGATGCCCTTTCTCCAGCACACAGACATTTGGGGACCTTGGAAATATTCG 1403
|||||
Db 241 CCTGAGGCGATGCCCTTTCTCCAGCACACAGACATTTGGGGACCTTGGAAATATTCG 300
QY 1404 TTCCAGGCTCTCTGTTCTCTGGACTTCAGATCTCTGGGGAGCCCCCTCCCTCCCTGAATCC 1463
|||||
Db 301 TTCCAGGCTCTCTGTTCTCTGGACTTCAGATCTCTGGGGAGCCCCCTCCCTCCCTGAATCC 359
QY 1464 CTGGCTTAGCTACCTTCCTCTGCTGCTGACCTTAAACACCTCAGGTCAAGAAAGA 1523
|||||
Db 360 CTGGCTTAGCTACCTTCCTCTGCTGCTGACCTTAAACACCTCAGGTCAAGAAAGA 419
QY 1524 GTTTGTTTATTTTAAATGGAGTCTCGTTCTGTCGCCCGAGGCTGAGTGCAGTA 1583
|||||
Db 420 GTTAGTTC-TATTTTAAATGGAGTCTCGTTCTGTCGCCCGAGGCTGAGTGCAGTA 478
QY 1584 GTGCAATCTCCGCTCAGTACAACTCCCTCCCTGCGGCTCAAGCGATCTCCACCTCA 1643
|||||
Db 479 GTGCAATCTCCGCTCAGTACAACTCCCTCCCTGCGGCTCAAGCGATCTCCCACTCA 538
QY 1644 GCGCGGAAGTAGTGGGACTATAGTGTGTACCATCACACCTGGCTAATTTTGTATT 1703
|||||
Db 539 GCGCGGAAGTAGTGGGACTATAGTGTGTACCATCACACCTGGCTAATTTTGTATT-TA 597
QY 1704 TTTGTAGACACAGGGTTTCGCGCATGTTGCCAGGCTGTTGAATTCCTGAGCTCAAGC 1763
|||||
Db 598 AAGGTAGACACAGGGTTTCGCGCATGTTGCCAGGCTGTTGTAGTTCCTGAGCTCAAGC 657
QY 1764 AACCTGCCCGGCTCGGGCTCCCAAAAGTACTGGGATTACAGCAG-AAGGCACCATGCCCA 1822
|||||
Db 658 AA-CTGCGGACTCGGGCTCCCAAAAGTACTGGGATGACCCCGGAAAGGCACATAGCCAG 716
QY 1823 GCGTAGATGTCTTATCCCAATCCTTTGGCAGGATGCGAGTCCA 1868
|||||
Db 717 GCGTAGATGTCTTATCCCAATCCTTTGGCAGGATGCGAGTCCA 762
```

```
RESULT 12
LOCUS BQ183276/c
DEFINITION UI-H-EUO-azr-e-05-0-UI.s1 NCI_CGAP_Carl Homo sapiens cDNA clone
BQ183276
linear EST 30-APR-2002
```

IMAGE: 5852524 3', mRNA sequence.  
 BQ183276  
 VERSION BQ183276.1 GI:20358826  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 629)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Seq primer: M13 FORWARD  
 POLYA-Yes

FEATURES  
 source  
 1..629  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NCI-CGAP\_Carl"  
 /tissue\_type="Osteoarthritic Cartilage"  
 /dev\_stage="Adult"  
 /lab\_host="PH10B (Life Technologies)"  
 /notes="Organ: Knee; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI-CGAP\_Carl is a cDNA library containing the following  
 tissue(s): Osteoarthritic Cartilage The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 TGATCAGCT.  
 TAG\_L1B=UI-H-EUO  
 TAG\_TISSUE=osteoarthritic cartilage  
 TAG\_SEQ=TGATCAGCT"

BASE COUNT 150 a 162 c 156 g 160 t 1 others  
 ORIGIN  
 Query Match 23.4%; Score 626.4; DB 14; Length 629;  
 Best Local Similarity 99.7%; Pred. No. 2.1e-60;  
 Matches 627; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2026 GACATGTATCCCTTTGACCCCAAGCTCTGCTCCTCCCTGACCAACCATGCCCTTTCCCTT 2085  
 |||||||  
 Db 629 GACATGTATCCCTTTGACCCCAAGCTCTGCTCCTCCCTGACCAACCATGCCCTTTCCCTT 570  
 |||||||

Qy 2086 TAACCTTCTCAACAGATACACAGGCCCTAAACTGCTTTACCTCCCTCCTACTAGTCAGG 2145  
 |||||||  
 Db 569 TAACCTTCTCAACAGATACACAGGCCCTAAACTGCTTTACCTCCCTCCTACTAGTCAGG 510  
 |||||||

Qy 2146 TTAGTGTGTGGAGTGACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAG 2205  
 |||||||  
 Db 509 TTAGTGTGTGGAGTGACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAG 450  
 |||||||

Qy 2206 TCATGTGGGTATGTCTGGGGTAGAGAGGGGTAGCAAGTTCAATGTCTCTCTTGGTCA 2265  
 |||||||  
 Db 449 TCATGTGGGTATGTCTGGGGTAGAGAGGGGTAGCAAGTTCAATGTCTCTCTTGGTCA 390  
 |||||||

Qy 2266 CATATCTCCCAAGCTCTGATCCCTGCCATGGGAGTGGACAGAAACATGAGGTCA 2325  
 |||||||

Db 389 CATATCTCCCAAGCTCCGATCCCTGCGCATGGGAAGTGGACAGGAACATGAGGTCA 330  
 |||||||

Qy 2326 CCTGCAGGCATCTTTACTGCAGCTCTGCCGCTCTGAGGGGAGAGGGGAGGAAGT 2385  
 |||||||

Db 329 CTTGCAGGCATCTTTACTGCAGCTCTGCCGCTCTGAGGGGAGAGGGGAGGAAGT 270  
 |||||||

Qy 2386 ATGCCGTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGTCTGAG 2445  
 |||||||

Db 269 ATGCCGTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGTCTGAG 210  
 |||||||

Qy 2446 CAGTACAGCGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTCGCTCCCTCAGGCAGA 2505  
 |||||||

Db 209 CAGTACAGCGCTCCAGTTTGGGCCCGATAAGGAAGTTCTCCGTCGCTCCCTCAGGCAGA 150  
 |||||||

Qy 2506 CGAGGAGGAGGCTGCATTTGCCAGTCTCTTCTGGGGCCCAAGGCGAGGTTGCAGGAGATC 2565  
 |||||||

Db 149 CGAGGAGGAGGCTGCATTTGCCAGTCTCTTCTGGGGCCCAAGGCGAGGTTGCAGGAGATC 90  
 |||||||

Qy 2566 CAATCCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAACCTGCAGTA 2625  
 |||||||

Db 89 CAATCCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAACCTGCAGTA 30  
 |||||||

Qy 2626 TTTTGAAGCAAAAAA 2654  
 |||||||

Db 29 TTTTGAAGCAAAAAA 1  
 |||||||

RESULT 13  
 AF188522 607 bp mRNA linear EST 02-MAR-2000  
 LOCUS AF188522 Homo sapiens ATCC HTB-12; SW1088 Homo sapiens cDNA clone  
 DEFINITION ISG 3, mRNA sequence.  
 ACCESSION AF188522  
 VERSION AF188522.1 GI:7144568  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 607)  
 Ye, Z. and Connor, J. R.  
 Identification of Iron Regulated Genes by Rescreening cDNA  
 Libraries from SSH with Antisense Probe from Three Iron Conditions  
 Unpublished (2000)  
 CONTACT: Ye Z  
 Neuroscience and Anatomy  
 Pennsylvania State University College of Medicine  
 500 University Drive, Hershey, PA 17033, USA  
 library screened by SSH and reverse Northern blot; increased  
 expression in iron loading was confirmed by Northern blot.

FEATURES  
 source  
 1..607  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="ISG 3"  
 /tissue\_type="astrocytoma"  
 /cell\_line="ATCC HTB-12; SW1088"  
 /note="Organ: liver"

BASE COUNT 154 a 153 c 155 g 145 t  
 ORIGIN  
 Query Match 22.6%; Score 605.4; DB 9; Length 607;  
 Best Local Similarity 99.8%; Pred. No. 4.4e-58;  
 Matches 606; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1794 GGGATTACCGCAGAGGCGACCATGCCAGGCTAGATGTCTTATCCCAATCCTTTGGC 1853  
 |||||||

Db 1 GGGATTACCGCAGAGGCGACCATGCCAGGCTAGATGTCTTATCCCAATCCTTTGGC 60  
 |||||||

Qy 1854 AGCATCAGCTCCACAGGCGATTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGTT 1913  
 |||||||

Db 61 AGGCATGACAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTATTAGCCCTCCCTGGGTT 120

QY 1914 AAGAGCCAGATAAGAGAAATCCCTTCTAGGTTTGGAAATGTTGTGTAAGAAAGAG 1973

Db 121 AAGAGCCAGATAAGAGAAATCCCTTCTAGGTTTGGAAATGTTGTGTAAGAAAGAG 180

QY 1974 AATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAGCAAACTCCCTGACATGA 2033

Db 181 AATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAGCAAACTCCCTGACATGA 240

QY 2034 TCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCACCATGCCCTTCTTAACTTCT 2093

Db 241 TCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCACCATGCCCTTCTTAACTTCT 300

QY 2094 CAAACAGATACCAGGCGCTAAACTGCTTTACCTCCCTCTACTAGTCAGTCAGTTAGGTGG 2153

Db 301 CAAACAGATACCAGGCGCTAAACTGCTTTACCTCCCTCTACTAGTCAGTCAGTTAGGTGG 360

QY 2154 TGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGATGATGATGATGATGATG 2213

Db 361 TGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGATGATGATGATGATGATG 420

QY 2214 GTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCCTCTGTCACATATCTC 2273

Db 421 GTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCCTCTGTCACATATCTC 480

QY 2274 CCAAGCTCTGATCCTGTCATGGGAAGTGGACAGAAACATGAGTCAATGACCTGCAGG 2333

Db 481 CCAAGCTCTGATCCTGTCATGGGAAGTGGACAGAAACATGAGTCAATGACCTGCAGG 540

QY 2334 CATCTTTACTGAGCTCTGCGGCTGGAGGGGAGAGGGGAGGAGGAAGATGCGCTG 2393

Db 541 CATCTTTACTGAGCTCTGCGGCTGGAGGGGAGAGGGGAGGAGGAAGATGCGCTG 600

QY 2394 CACATTT 2400

Db 601 CACATTT 607

RESULT 14

LOCUS BI767046

DEFINITION 603054267f1 NIH\_MGC\_122 Homo sapiens cdna clone IMAGE:5203600 5', mRNA linear EST 25-SEP-2001

ACCESSION BI767046

VERSION BI767046.1 GI:15758624

KEYWORDS EST

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 740)

TITLE NIH-MGC http://mgc.ncl.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL1510 row: f column: 17  
High quality sequence stop: 707.  
Location/Qualifiers 1..740

FEATURES source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5203600"  
/clone\_lib="NIH\_MGC\_122"  
/lab\_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

BASE COUNT 171 a 198 c 222 g 149 t

ORIGIN

Query Match 22.5%; Score 601.6; DB 13; Length 740;  
Best Local Similarity 96.2%; Pred. No. 9.8e-58;  
Matches 712; Conservative 0; Mismatches 19; Indels 9; Gaps 9;

QY 42 ATGAGGGGACGCCAGGCGCGCTTGGCCACCCGGGATGGGAGGCCGCCAGGCGCTCCCG 101

Db 1 ATGAGGGGACGCCAGGCGCGCTTGGCCACCCGGGATGGGAGGCCGCCAGGCGCTCCCG 60

QY 102 CCCTGCAGTCC-GAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCT 160

Db 61 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCT 120

QY 161 CGGCCAAACATGTTCTCTGATCCAAATTCAGACACGGGCGCTTCTCTCCGGAACTTCAT 220

Db 121 CGGCCAAACATGTTCTCTGATCCAAATTCAGACACGGGCGCTTCTCTCCGGAACTTCAT 180

QY 221 AGCCACCCGTCGGCAGTAGACTTCAGGAACAGGTGGTGTGCTGTGGATGGCGTGAGAGTAA 280

Db 181 AGCCACCCGTCGGCAGTAGACTTCAGGAACAGGTGGTGTGCTGTGGATGGCGTGAGAGTAA 240

QY 281 GCTGCAGATCTGGGACACCGCTGGGCGAGCAACGGTTCCGAAGCGTCACCATGCTTATTA 340

Db 241 GCTGCAGATCTGGGACACCGCTGGGCGAGCAACGGTTCCGAAGCGTCACCATGCTTATTA 300

QY 341 CAGAGATGCTCAGGCGCTTCTCTGCTGTATGACATCACCAACAAATCTCTTTCGACAA 400

Db 301 CAGAGATGCTCAGGCGCTTCTCTGCTGTATGACATCACCAACAAATCTCTTTCGACAA 360

QY 401 CATCAGGCGCTGGCTCAGTACATTCATGATGATGCCAGAGAGGAGCTGG-TGATCATGC 459

Db 361 CATCAGGCGCTGGCTCAGTACATTCATGATGATGCCAGAGAGGAGCTGGTGTGATCATGC 420

QY 460 TGCTAGGCAACAGGCGGATATGAGCAGCGAAGAGTATCCGTTCCGAAGAGCGGAGAGA 519

Db 421 TGCTAGGCAACAGGCGGATATGAGCAGCGAAGAGTATCCGTTCCGAAGAGCGGAGAGA 480

QY 520 CTTTGGCCAGGAGTACGGTGTTCCTTCTCTTCCCTTCCCT-GGAGACACCGGCCCAAGAC-TGGCATGAA 577

Db 481 CTTTGGCCAGGAGTACGGTGTTCCTTCTCTTCCCTTCCCTGGGAGAACAGTGCCTTGGCATGAA 540

QY 578 TGTGGAGTTA-GCCTTTTCTGGCCATGCGCAAGAACTGAAATACCGGGCGGCGCATCAGG 636

Db 541 TGTGGAGTTAGGCTTTTCTTGGCATGCGCAGGGAACCTTGAATACCGGGCGGCGCATCAGG 600

QY 637 CGGATGAGCCCGAGCTTCCAGA-TCCGAGACTATGTAG-TCCCAAGAAAGCGCTCCAG 694

Db 601 CGGATGAGCCCGAGCTTCCAGATTCGAGACTATGTAGAAAGTCCCAAGAAAGCGCTCCAG 660

QY 695 C-TGCTGTCTCTCATGTGAA-TCCGAGGGGCGAGAGGAGGCTCTGGAGGCGACACAGG 752

Db 661 CTTGTGTCTCTCATGTGAA-TCCGAGGGGCGAGAGGAGGCTCTGGAGGCGACACAGG 720

QY 753 ATGCAGCCTTCCCGCTCCCA 772

Db 721 ATGCAGCCTTCTCTCCATCCA 740

RESULT 15

LOCUS BG253976

691 bp mRNA linear EST 13-FEB-2001

DEFINITION	602366910F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4475067 5', mRNA sequence.									
ACCESSION	BG253976									
VERSION	BG253976.1 GI:12763792									
KEYWORDS	EST.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 (bases 1 to 691).									
TITLE	NIH-MGC http://mgc.nci.nih.gov/.									
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)									
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10300 row: c column: 04 High quality sequence stop: 650.									
FEATURES	Location/Qualifiers									
source	1..691									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/clone="IMAGE:4475067"									
	/clone_lib="NIH_MGC_91"									
	/lab_host="DH10B (phage-resistant)"									
	/tissue_type="adenocarcinoma, cell line"									
	/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."									
BASE COUNT	175 a 185 c 190 g 141 t									
ORIGIN										
Query Match	22.5%; Score 600.6; DB 12; Length 691;									
Best Local Similarity	97.6%; Pred. No. 1.3e-57;									
Matches 663; Conservative	0; Mismatches 9; Indels 7; Gaps 5;									
QY	287	GATCTGGGACACCGCTGGCGAGGAACGGTTCGGAAGCGGTCA	CCCCATGCCCATGCTTATTACAGAGA	346						
Db	14	GATCTGGGACACCGCTGGCGAGGAACGGTTCGGAAGCGGTCA	CCCCATGCCCATGCTTATTACAGAGA	73						
QY	347	TGCTCAGGCGCTTGCTTCTGCTGTATGATACATCACCAACAAATCTCTTCGACAAACATCAG	406							
Db	74	TGCTCAGGCGCTTGCTTCTGCTGTATGATACATCACCAACAAATCTCTTCGACAAACATCAG	133							
QY	407	GGCTTGCTCACTGAGATTCATGAGTATGCCAGGAGGACGTGTGTATCATGCTGCTAGG	466							
Db	134	GGCTTGCTCACTGAGATTCATGAGTATGCCAGGAGGACGTGTGTATCATGCTGCTAGG	193							
QY	467	CACAAGCGGATATGACGACGAAAGAGTGATCCGTTCCGAACGAGGAGACCTTGGC	526							
Db	194	CACAAGCGGATATGACGACGAAAGAGTGATCCGTTCCGAACGAGGAGACCTTGGC	253							
QY	527	CAGGAGTACGGTGTTCCTTCCTTGGAGACGACGCCAAGACTGGCATGAATGTGGAGTT	586							
Db	254	CAGGAGTACGGTGTTCCTTCCTTGGAGACGACGCCAAGACTGGCATGAATGTGGAGTT	313							
QY	587	AGCCTTCTGGCCATGCCCAAGGAACCTGAATACCGGCGCGGCGCATCAGGCGGATGAGCC	646							
Db	314	AGCCTTCTGGCCATGCCCAAGGAACCTGAATACCGGCGCGGCGCATCAGGCGGATGAGCC	373							
QY	647	CAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGCTT	706							
Db	374	CAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGCTT	433							
QY	707	CATGTGAATCCCGGGGCGAGAGGAGGCGTCTTGGAGGCACACAGGATGCAGCGCTTCCG	766							



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 13:13:23 ; Search time 4598 Seconds  
(without alignments)  
16924.946 Million cell updates/sec

Title: US-09-817-199a-1  
Perfect score: 2674  
Sequence: 1 ttccctgcgcggcgccact.....aaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2161.8	80.8	2195	9	AK057069	AK057069 Homo sapi
2	1562	73.4	1977	9	AK098068	AK098068 Homo sapi
3	1354.8	50.7	3105	9	AK054846	AK054846 Homo sapi
4	1086	40.6	1116	6	AX236082	AX236082 Sequence
5	938.4	35.1	1700	9	BC016615	BC016615 Homo sapi
6	576	21.5	576	6	AX236084	AX236084 Sequence
7	563.6	21.1	690	10	AF233582	AF233582 Mus muscu
8	315.6	11.8	1513	9	BC007681	BC007681 Homo sapi
9	312.4	11.7	1098	10	RNU18771	U18771 Rattus norv
10	304.6	11.4	1320	9	AB027137	AB027137 Homo sapi
11	303.2	11.3	573	9	AF498952	AF498952 Homo sapi
12	292.6	10.9	1340	6	AR062279	AR062279 Sequence
13	234.2	8.8	2536	3	AY061826	AY061826 Drosophil
14	214.8	8.0	127587	9	HS941P9	Z95331 Human DNA s
15	214.2	8.0	102332	9	AC010352	AC010352 Homo sapi
16	214.2	8.0	107465	9	AC010434	AC010434 Homo sapi
17	212.8	8.0	146740	9	CNS01D7X	AL132989 Human chr
18	212.6	8.0	174309	9	AC016656	AC016656 Homo sapi
19	212.6	8.0	177640	9	AC016652	AC016652 Homo sapi
20	211.6	7.9	148841	9	AC011462	AC011462 Homo sapi
21	210.4	7.9	199601	9	AL389915	AL389915 Human DNA
22	210.2	7.9	90000	9	AC117467	AC117467 Homo sapi
23	209.6	7.8	141771	9	AC091821	AC091821 Homo sapi
24	209	7.8	110000	2	AC013391_0	AC013391 Homo sapi
25	209	7.8	125980	9	AC093223	AC093223 Homo sapi
26	209	7.8	157141	9	AC011470	AC011470 Homo sapi
27	209	7.8	176304	9	AC013787	AC013787 Homo sapi
28	208.4	7.8	167509	2	AC012107	AC012107 Homo sapi
29	208	7.8	43790	9	AC113388	AC113388 Homo sapi
30	208	7.8	86685	9	AC118457	AC118457 Homo sapi
31	208	7.8	107995	2	HSJ356711	AL121971 Homo sapi
32	208	7.8	172048	9	HS179N16	Z95152 Homo sapien
33	207.6	7.8	132721	9	AC004552	AC004552 Homo sapi
34	207.6	7.8	207366	9	AC090001	AC090001 Homo sapi
35	207.4	7.8	136222	9	AC004099	AC004099 Homo sapi
36	207.2	7.7	63905	9	AC022515	AC022515 Homo sapi
37	207.2	7.7	186330	9	AC007376	AC007376 Homo sapi
38	207.2	7.7	271814	2	AC027406	AC027406 Homo sapi
39	207	7.7	154289	2	AC027714	AC027714 Homo sapi
40	207	7.7	161757	2	AC044901	AC044901 Homo sapi
41	207	7.7	163528	2	AC130416	AC130416 Homo sapi
42	207	7.7	187399	9	AC090970	AC090970 Homo sapi
43	207	7.7	209797	2	AC005748	AC005748 Homo sapi
44	207	7.7	346940	2	AC107420	AC107420 Homo sapi
45	206.8	7.7	157113	9	AC083797	AC083797 Homo sapi

ALIGNMENTS

RESULT 1  
AK057069  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ32507 fis, clone SMINT1000048, moderately similar to Mus musculus GTPase Rab37 mRNA.  
ACCESSION AK057069  
VERSION AK057069.1 GI:16552648  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens small intestine cDNA to mRNA, clone\_lib:SMINT1 clone:SMINT1000048.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,

Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,  
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,  
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.  
and Isogai,T.

## NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 2195)

Isogai,T., Otsuki,T. and Sugiyama,T.

## Direct Submission

Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.

## FEATURES

## Source

Location/Qualifiers  
1..2195  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SMINT1000048"  
/tissue.type="small intestine"  
/clone\_lib="SMINT1"  
/note="cloning vector: pWE185FL3"

BASE COUNT 511 a 596 c 582 g 506 t

## ORIGIN

Query Match 80.8%; Score 2161.8; DB 9; Length 2195;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2163; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 472 AGCGGATATAGCAGCGAAGAGTAGTCGGTTCGGAAGCGGAGAGACCTTGCGCAGGG 531  
DB 31 AGCGGATATAGCAGCGAAGAGTAGTCGGTTCGGAAGCGGAGAGACCTTGCGCAGGG 90

QY 532 AGTACGGTGTTCCTTCCCTGGAGCACCAGCCCAAGACTGGCATGAATGTGGAGTTAGCCCT 591  
DB 91 AGTACGGTGTTCCTTCCCTGGAGCACCAGCCCAAGACTGGCATGAATGTGGAGTTAGCCCT 150

QY 592 TTCTGGCCATCGCCAGGAAGTGAATATACCGGGCCGGCATGAGGGGATGAGCCGACGT 651  
DB 151 TTCTGGCCATCGCCAGGAAGTGAATATACCGGGCCGGCATGAGGGGATGAGCCGACGT 210

QY 652 TCCAGATCCGAGACTATGTAGAGTCCAGAAAGAGGCTCCAGCTGCTGCTCTTCATGT 711  
DB 211 TCCAGATCCGAGACTATGTAGAGTCCAGAAAGAGCGCTCCAGCTGCTGCTCTTCATGT 270

QY 712 GAATCCAGGGGGCAGAGAGGAGGCTCTGGAGCACCAGAGATGCAGCTTCCCTCCC 771  
DB 271 GAATCCAGGGGGCAGAGAGGAGGCTCTGGAGCACCAGAGATGCAGCTTCCCTCCC 330

QY 772 AGGCTTGCTTATTCGAAGAGGCTGAGCCCAATGGGAGAAAGATGGAGGACTCACTGCAC 831  
DB 331 AGGCTTGCTTATTCGAAGAGGCTGAGCCCAATGGGAGAAAGATGGAGGACTCACTGCAC 390

QY 832 AGCGCTTCCCTAGCAGGAGGTATACATCCCACTCTACTTACTTCTGCTGGGCTCCCGG 891  
DB 391 AGCGCTTCCCTAGCAGGAGGTATACATCCCACTCTACTTACTTCTGCTGGGCTCCCGG 450

QY 892 CATCCAGGGAGGTTAAACACATTTAGCTTTTATTTTAAATAGTACATAATTTAATACCAA 951  
DB 451 CATCCAGGGAGGTTAAACACATTTAGCTTTTATTTTAAATAGTACATAATTTAATACCAA 510

QY 952 AAAAGGGCCCTGGATCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCCTTTCTA 1011  
DB 511 AAAAGGGCCCTGGATCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCCTTTCTA 570

QY 1012 GGACTTGGGGGGCGGCGCTCCCTTAAGCATACAAAGCTGGTGTGCTCCAGCTCAG 1071  
DB 571 GGACTTGGGGGGCGGCGCTCCCTTAAGCATACAAAGCTGGTGTGCTCCAGCTCAG 630

QY 1072 CCCAGGGGACACACATGACATCTTTGGGGGTGAGGCGAGGTAACTACTCCATCGCACCCCTC 1131  
DB 631 CCCAGGGGACACACATGACATCTTTGGGGGTGAGGCGAGGTAACTACTCCATCGCACCCCTC 690

QY 1132 AGTTCAGCTGGACAGAGGCTCAGGTGACCCCGACGCTTCACTGTCTCCGCTCTCCAGAG 1191  
DB 691 AGTTCAGCTGGACAGAGGCTCAGGTGACCCCGACGCTTCACTGTCTCCGCTCTCCAGAG 750

QY 1192 CTTATCTTCGCCCCATCTCCCAATAAAGTGGGCCCTTGTGCTGAGGAAGACCAAGCC 1251  
DB 751 CTTATCTTCGCCCCATCTCCCAATAAAGTGGGCCCTTGTGCTGAGGAAGACCAAGCC 810

QY 1252 TCAGGGAAGATAAGAGATATGAGATGGAGGGGAGGACAAAGGCGCAGAGTAGGGTC 1311  
DB 811 TCAGGGAAGATAAGAGATATGAGATGGAGGGGAGGACAAAGGCGCAGAGTAGGGTC 870

QY 1312 TAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCTTTTCTCCAGCA 1371  
DB 871 TAGCTGGCTATCTCTGGCCTTACTAACACCCCCCTGGAGGCATGCCCTTTTCTCCAGCA 930

QY 1372 CACAAGCACATTTGGGCACCTGGAAATATTTGGTTCAGGCTCCTGTTCTCTGGACTTCAG 1431  
DB 931 CACAAGCACATTTGGGCACCTGGAAATATTTGGTTCAGGCTCCTGTTCTCTGGACTTCAG 990

QY 1432 ATCTTGGGGAGCCCTCCCGCCCTGAAATCCCTGGCTTAGCTTACCTTCCCTCGCTGCA 1491  
DB 991 ATCTTGGGGAGCCCTCCCGCCCTGAAATCCCTGGCTTAGCTTACCTTCCCTCGCTGCA 1050

QY 1492 CTTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTGTGTTTATTTTGTGAAATGGAG 1551  
DB 1051 CTTAAAAACCTCAGGTCAGAACTAGGAAAAGAGTTTGTGTTTATTTTGTGAAATGGAG 1110

QY 1552 TCTCGTCTCTGCCAGGCTCAGGTGCGTAGTGCATCTCCGCTCACTACAACCTCCA 1611  
DB 1111 TCTCGTCTCTGCCAGGCTCAGGTGCGTAGTGCATCTCCGCTCACTACAACCTCCA 1170

QY 1612 CTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTG 1671  
DB 1171 CTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCGCCGAAGTAGCTGGGACTATAGGTG 1230

QY 1672 TGTACCATCACAGCTGGCTAAATTTTGTATTTTGTAGACACAGGTTTCGCCATGTTG 1731  
DB 1231 TGTACCATCACAGCTGGCTAAATTTTGTATTTTGTAGACACAGGTTTCGCCATGTTG 1290

QY 1732 CCCAGGCTGCTTGAATTCCTGAGCTCAAGCAACCTCGCGCTCGGCTCCCAAGTA 1791  
DB 1291 CCCAGGCTGCTTGAATTCCTGAGCTCAAGCAACCTCGCGCTCGGCTCCCAAGTA 1350

QY 1792 CTGGGATACACGAGAAGGCACCATGCCAGGCTAGATGTGCTTATCCCAATCCTTTG 1851  
DB 1351 CTGGGATACACGAGAAGGCACCATGCCAGGCTAGATGTGCTTATCCCAATCCTTTG 1410

QY 1852 GCAGGATCAGCTCCACAGGGGATTTCTTCAAGAGCTGAAGTGTTCAGCCCTCTGCGG 1911  
DB 1411 GCAGGATCAGCTCCACAGGGGATTTCTTCAAGAGCTGAAGTGTTCAGCCCTCTGCGG 1470

QY 1912 TTAAGAGCCAGATAAGGAGAAATCCCTTTCTAGGTTTGGAAATGTGTTGTGAAAAAAG 1971  
DB 1471 TTAAGAGCCAGATAAGGAGAAATCCCTTTCTAGGTTTGGAAATGTGTTGTGAAAAAAG 1530

QY 1972 AGAAATCCCTGGCTCTCGAGCTGGGGAGACAAGATAAGCAAACTCCCTCGCATG 2031  
DB 1531 AGAAATCCCTGGCTCTCGAGCTGGGGAGACAAGATAAGCAAACTCCCTCGCATG 1590

QY 2032 TATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCCCACTGCTTCTTAACTT 2091  
DB 1591 TATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCCCACTGCTTCTTAACTT 1650

QY 2092 CTCAAAACAGATACCGGGCTAAACTGCTTTACCTCCCTCTACTGAGTCAGGTAGGT 2151

```
|||||
Db 1651 CTCAACAGATACACAGGCGCTAACTGCTTTACCTCCCTCTCTACTGAGTCAGGTAGGT 1710
QY 2152 GGTGGAGGTACCATTCCTCGAGTTAAACCAATGCAATATGAGTAAACAAGTCATGT 2211
Db 1711 GGTGGAGGTACCATTCCTCGAGTTAAACCAATGCAATATGAGTAAACAAGTCATGT 1770
QY 2212 GGTATGTCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTCTCTCTGTTGTCACATATC 2271
Db 1771 GGTATGTCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTCTCTCTGTTGTCACATATC 1830
QY 2272 TCCCAAGCTCTGATCCCTGTCATGGGAAGTGGACAGGAACATGAGTTCATGACCTGCA 2331
Db 1831 TCCCAAGCTCCGATCCCTGTCATGGGAAGTGGACAGGAACATGAGTTCATGACCTGCA 1890
QY 2332 GGCATCTTTACTGTCAGCTCTGCCGCTCTGGAGGGGAGGGGAGGAGGAAGTATGCGC 2391
Db 1891 GGCATCTTTACTGTCAGCTCTGCCGCTCTGGAGGGGAGGGGAGGAGGAAGTATGCGC 1950
QY 2392 TGCACATTTCTGAGGCTACTGCTATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCA 2451
Db 1951 TGCACATTTCTGAGGCTACTGCTATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCA 2010
QY 2452 GCGGCTCCAGTTTGGGCCCCGATTAAGAAAGTTCTCCGTGGCCCTCCCTCAGGCAGAGCAGGG 2511
Db 2011 GCGGCTCCAGTTTGGGCCCCGATTAAGAAAGTTCTCCGTGGCCCTCCCTCAGGCAGAGCAGGG 2070
QY 2512 AGGAGGCTGACATTTCCAGTCTCTTCTGGGGCCCAAGCAGGTTGCAGGAGATCAATCC 2571
Db 2071 AGGAGGCTGACATTTCCAGTCTCTTCTGGGGCCCAAGCAGGTTGCAGGAGATCAATCC 2130
QY 2572 CATGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAAATTAACCTGACGATTTTGG 2631
Db 2131 CATGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAAATTAACCTGACGATTTGG 2190
QY 2632 AAAGC 2636
|||||
Db 2191 AAAGC 2195

RESULT 2
AK098068
LOCUS
DEFINITION Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION AK098068
VERSION AK098068.1 GI:21758001
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens trachea cDNA to mRNA, clone_lib:TRACH2
clone:TRACH2000540.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A.,
Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K.,
Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1977)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
```

```
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. 1977
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TRACH2000540"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="cloning vector: pME18SFL3"
57..728
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC05227.1"
/db_xref="gi:21758002"
/translation="MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVKTKT
FLIOFKDGAFLSGTFIATVGIDFRNKVVTVDGVKVLQIWDAGDERFSVTHAYRD
AQALLLLYDIYTKSSFDNIRAWLFEIHEYAQRDVIYLLGNKADMSRVSRSDEGET
LAREYGVFFLETSARTGMNVELAFIALAKELKRYAGHQADEPFSQIIRDYVESKRRSS
CCSPM"
BASE COUNT 442 a 560 c 539 g 436 t
ORIGIN
Query Match 73.4%; Score 1962; DB 9; Length 1977;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1962; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCGGCTCGCGGCGGCACTGCTCACCTCTCGTCAGGGACATGACGGGCACGCCAGGGC 60
Db 16 TTCGGCTCGCGGCGGCACTGCTCACCTCTCGTCAGGGACATGACGGGCACGCCAGGGC 75
QY 61 CCGTTGCCAACCAGGAGTGGCGAGGCCCCCGAGCGTCCCGCCCTCGAGTCCGAGTAGC 120
Db 76 CCGTTGCCAACCAGGAGTGGCGAGGCCCCCGAGCGTCCCGCCCTCGAGTCCGAGTAGC 135
QY 121 ACCTCAGCGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTCCCTCA 180
Db 136 ACCTCAGCGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTCCCTCA 195
QY 181 TCCCAATTCAGACAGCGGCGCTTCTCTCGCGAACCTTCATAGCCACCGTCGGCATAGACT 240
Db 196 TCCCAATTCAGACAGCGGCGCTTCTCTCGCGAACCTTCATAGCCACCGTCGGCATAGACT 255
QY 241 TCAGGAACAAGGTGGTGAAGTGTGGATGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 300
Db 256 TCAGGAACAAGGTGGTGAAGTGTGGATGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 315
QY 301 CTGGCAGCAAGCGTTCCGAAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGTTGC 360
Db 316 CTGGCAGCAAGCGTTCCGAAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGTTGC 375
QY 361 TTCTGCTGTATGACATCAGCAACAAATCTCTTTTCGACAACTCAGGCGCTGGCTCAGTGC 420
Db 376 TTCTGCTGTATGACATCAGCAACAAATCTCTTTTCGACAACTCAGGCGCTGGCTCAGTGC 435
QY 421 AGATTTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTGTAGTGAACAACAAGCGGATA 480
Db 436 AGATTTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTGTAGTGAACAACAAGCGGATA 495
QY 481 TGAGCAGCAAGAGATGATCCGTTCCGAAAGCGGAGACCTTGGCCAGGGAGTAGCGGTG 540
Db 496 TGAGCAGCAAGAGATGATCCGTTCCGAAAGCGGAGACCTTGGCCAGGGAGTAGCGGTG 555
QY 541 TTCCCTTCTGAGAGACCGAGCGCAAGACTGGCATGAATGTGGAGTTAGCTTTCTGGCCA 600
Db 556 TTCCCTTCTGAGAGACCGAGCGCAAGACTGGCATGAATGTGGAGTTAGCTTTCTGGCCA 615
QY 601 TCGCCAAAGAACTGAATAACCGGGCGGCATCAGCGGATGAGGCCAGCTTCCAGATCC 660
|||||
```

Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.

#### FEATURES

##### source

1. 1977  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="TRACH2000540"  
/tissue\_type="trachea"  
/clone\_lib="TRACH2"  
/note="cloning vector: pME18SFL3"  
57..728  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAC05227.1"  
/db\_xref="gi:21758002"  
/translation="MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVKTKT  
FLIOFKDGAFLSGTFIATVGIDFRNKVVTVDGVKVLQIWDAGDERFSVTHAYRD  
AQALLLLYDIYTKSSFDNIRAWLFEIHEYAQRDVIYLLGNKADMSRVSRSDEGET  
LAREYGVFFLETSARTGMNVELAFIALAKELKRYAGHQADEPFSQIIRDYVESKRRSS  
CCSPM"

##### CDS

Db 616 TCGCCAAGGAACGTAATAACCGCGCGGCATCAGCGGATGAGCCAGCTCCAGATCC 675  
QY 661 GAGACTATGTAGAGTCCCAAGAAAGCGTCCAGCTGCTGCTCTTTCATGTGAATCCCAAG 720  
Db 676 GAGACTATGTAGAGTCCCAAGAAAGCGTCCAGCTGCTGCTCTTTCATGTGAATCCCAAG 735  
QY 721 GGGGCGAGAGGAGGCTCTGGAGGCACACAGGATGAGGCTTCCGCCCTCCAGGCGCTGGC 780  
Db 736 GGGGCGAGAGGAGGCTCTGGAGGCACACAGGATGAGGCTTCCGCCCTCCAGGCGCTGGC 795  
QY 781 TTATTCAGAGAGCTGAGCCATGGGAGAAAGATGGAGGACTCACTGCACAGCCCTTC 840  
Db 796 TTATTCAGAGAGCTGAGCCATGGGAGAAAGATGGAGGACTCACTGCACAGCCCTTC 855  
QY 841 CTAGCGGAGCTATCTCAACTCTCTACTTGTAGTTCCTGCGGTCTCCCGCATCCACAG 900  
Db 856 CTAGCGGAGCTATCTCAACTCTCTACTTGTAGTTCCTGCGGTCTCCCGCATCCACAG 915  
QY 901 GGAGGCTAAACACTTAGCTTTTATTTTAACTAGTACATAATTTAATACCAAAAAGGCGC 960  
Db 916 GGAGGCTAAACACTTAGCTTTTATTTTAACTAGTACATAATTTAATACCAAAAAGGCGC 975  
QY 961 CTGGATCCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGCCTTCTAGGACTTGG 1020  
Db 976 CTGGATCCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGCCTTCTAGGACTTGG 1035  
QY 1021 GGGCGGCGCTCCCTCTTAAGCATAAACAAAGTGGTGTCTCCAGCTCAGCCCGCAGGGG 1080  
Db 1036 GGGCGGCGCTCCCTCTTAAGCATAAACAAAGTGGTGTCTCCAGCTCAGCCCGCAGGGG 1095  
QY 1081 ACACAGATGACCTTGGGGGTGAGGAGGTAATGACTCCATCGCACCTTCAGTTCAGCT 1140  
Db 1096 ACACAGATGACCTTGGGGGTGAGGAGGTAATGACTCCATCGCACCTTCAGTTCAGCT 1155  
QY 1141 GGACAGAGGCTCAGGTACCCCGCTTCTACTCTCTCCCGCTCTCCAGGAGCTTACTTTC 1200  
Db 1156 GGACAGAGGCTCAGGTACCCCGCTTCTACTCTCTCCCGCTCTCCAGGAGCTTACTTTC 1215  
QY 1201 GCCCATCTCCCAATAAGTGGGCGCTTGTGCTGTAGGAAAGCAAGAGCTCAGGGAAG 1260  
Db 1216 GCCCATCTCCCAATAAGTGGGCGCTTGTGCTGTAGGAAAGCAAGAGCTCAGGGAAG 1275  
QY 1261 ATAAGAGATGAGATGGAGGGGAGGACAAAGGGCAGAGATGAGGTCTAGCTGGCT 1320  
Db 1276 ATAAGAGATGAGATGGAGGGGAGGACAAAGGGCAGAGATGAGGTCTAGCTGGCT 1335  
QY 1321 ATCTCTGCGCTTACTAACACCCCTTGAGGCGATGCCCTTTCTCAGCACACAAGCAC 1380  
Db 1336 ATCTCTGCGCTTACTAACACCCCTTGAGGCGATGCCCTTTCTCCAGCACACAAGCAC 1395  
QY 1381 ATTTGGGACCTGGAATATTTGGTTCCAGGCTCCTGTCTCTGAGCTTCAGATCTCTGGG 1440  
Db 1396 ATTTGGGACCTGGAATATTTGGTTCCAGGCTCCTGTCTCTGAGCTTCAGATCTCTGGG 1455  
QY 1441 GAGCCCTCCCGCCCTGATCCCTGGCTTAGCTACCTTCCCTGCTGTGCACCTAAAAC 1500  
Db 1456 GAGCCCTCCCGCCCTGATCCCTGGCTTAGCTACCTTCCCTGCTGTGCACCTAAAAC 1515  
QY 1501 CTCAGGTCAGAACTAGGAAAAGAGTTTGTATTTTATTTTGAATGGAGTCTCGTTCT 1560  
Db 1516 CTCAGGTCAGAACTAGGAAAAGAGTTTGTATTTTATTTTGAATGGAGTCTCGTTCT 1575  
QY 1561 GTCGCCCGAGGCTGAGGTGCAGTAGTGAATCTCCGCTCACTACAACTCCACCTCCCTGGG 1620  
Db 1576 GTCGCCCGAGGCTGAGGTGCAGTAGTGAATCTCCGCTCACTACAACTCCACCTCCCTGGG 1635  
QY 1621 GCTCAAGGATCTCCACCTCAGCCCGCAAGTAGTGGACTATAGGTGTGTACCATC 1680  
Db 1636 GCTCAAGGATCTCCACCTCAGCCCGCAAGTAGTGGACTATAGGTGTGTACCATC 1695  
QY 1681 ACACCTGGCTAATTTTGTATTTTGTAGACACAGGTTTTCGCTATGTTGCCAGGCTG 1740  
Db 1696 ACACCTGGCTAATTTTGTATTTTGTAGACACAGGTTTTCGCTATGTTGCCAGGCTG 1755

QY 1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCGCTCGGCCTCCCAAAGTACTGGGATTA 1800  
Db 1756 GTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCGCTCGGCCTCCCAAAGTACTGGGATTA 1815  
QY 1801 CAGGAGAGAGGACCATGCCAGGCTAGATGTCTTATATCCCAATCCTTTGGCAGGCAATG 1860  
Db 1816 CAGGAGAGAGGACCATGCCAGGCTAGATGTCTTATATCCCAATCCTTTGGCAGGCAATG 1875  
QY 1861 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTCTTAGCCCTCTCGGTTAAGAGCC 1920  
Db 1876 CAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTCTTAGCCCTCTCGGTTAAGAGCC 1935  
QY 1921 AGATAGGAGAAATCCCTTTCTCTAGTGTGGATGTGTTGTG 1962  
Db 1936 AGATAGGAGAAATCCCTTTCTCTAGTGTGGATGTGTTGTG 1977  
  
RESULT 3  
AK054846  
LOCUS  
DEFINITION  
Homo sapiens cDNA FLJ30284 fis, clone BRACE2002812, moderately similar to Mus musculus GTPase Rab37 (Rab37) mRNA.  
ACCESSION  
AK054846  
VERSION  
AK054846.1 GI:16549463  
KEYWORDS  
Homo sapiens cerebellum cDNA to mRNA, clone\_lib:BRACE2  
SOURCE  
clone:BRACE2002812.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
  
REFERENCE  
1  
AUTHORS  
Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yanashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fuji,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T., NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 3105)  
Isogai,T., Otsuki,T. and Sugiyama,T.  
Direct Submission  
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.  
  
FEATURES  
source  
i. .3105  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="BRACE2002812"  
/tissue.type="cerebellum"  
/clone\_lib="BRACE2"  
/note="cloning vector: pME18SFL3"  
BASE COUNT 680 a 905 c 898 g 622 t  
ORIGIN

Query Match 50.7%; Score 1354.8; DB 9; Length 3105;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1362; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
  
QY 589 CTTTCTGGCGATCGCAGGAACTGAAATACCGGCGCGGATCATGGCGGATGAGCCCA 648  
Db 1732 CCATTGCTCTTCTTCAAGGGAACCTGAAATACCGGCGCGGATCATGGCGGATGAGCCCA 1791

QY 649 GCTTCCAGATCGAGACATATGTAGAGTCCAGAAAGCGCTCCAGCTGCTGCTCCTTCA 708  
|||||  
Db 1792 GCTTCCAGATCGAGACATATGTAGAGTCCAGAAAGCGCTCCAGCTGCTGCTCCTTCA 1851  
QY 709 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGCACACAGGATGACAGCTTCCCCCT 768  
|||||  
Db 1852 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGCACACAGGATGACAGCTTCCCCCT 1911  
QY 769 CCCAGGCTGGCTTATCCAGAGGCTGAGCAATGGGGAGAAAGATGAGAGCTACAGT 828  
|||||  
Db 1912 CCCAGGCTGGCTTATCCAGAGGCTGAGCAATGGGGAGAAAGATGAGAGCTACAGT 1971  
QY 829 CACAGCGCTTCTAGCAGGAGCTATACCTCCAACCTCTACTTGAGTTCTCGGCTCC 888  
|||||  
Db 1972 CACAGCGCTTCTAGCAGGAGCTATACCTCCAACCTCTACTTGAGTTCTCGGCTCC 2031  
QY 889 CCGCATCCACAGGGAGGTAACACTTAGCTTTTATTTTAATAGTACATAATTTAATAC 948  
|||||  
Db 2032 CCGCATCCACAGGGAGGTAACACTTAGCTTTTATTTTAATAGTACATAATTTAATAC 2091  
QY 949 CAAAAAGCGCTGGATCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 1008  
|||||  
Db 2092 CAAAAAGCGCTGGATCCCAAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 2151  
QY 1009 CTAGGACTTTGGGGGCGGCCCTCCCTCCTAAGCATATAACAAAGGTGGTGTCTCCAGCT 1068  
|||||  
Db 2152 CTAGGACTTTGGGGGCGGCCCTCCCTCCTAAGCATATAACAAAGGTGGTGTCTCCAGCT 2211  
QY 1069 CAGCCCCAGGGGACACAGATGACATCTTTGGGGTGGGAGGTAATGACTCCATCGCAC 1128  
|||||  
Db 2212 CAGCCCCAGGGGACACAGATGACATCTTTGGGGTGGGAGGTAATGACTCCATCGCAC 2271  
QY 1129 CTCAGTTAGCTGGACAGAGGCTCAGGTGACCCAGCCCTCAGTGTCTCCCGCTCCAG 1188  
|||||  
Db 2272 CTCAGTTAGCTGGACAGAGGCTCAGGTGACCCAGCCCTCAGTGTCTCCCGCTCCAG 2331  
QY 1189 GAGCTTATCTCGCCCCATCTCCAAATAGTGGGCCCTTGTGCTGTAGGAAGACCAAA 1248  
|||||  
Db 2332 GAGCTTATCTCGCCCCATCTCCAAATAGTGGGCCCTTGTGCTGTAGGAAGACCAAA 2391  
QY 1249 GCCTCAGGGAAGATAGAGATATGAGATGGAGGGGGAGGACAAAGGGGCGAGAGTAGG 1308  
Db 2392 GCCTCAGGGAAGATAGAGATATGAGATGGAGGGGGAGGACAAAGGGGCGAGAGTAGG 2451  
QY 1309 GTCTAGCTGGCTATCTCGCCCTTACTAACCCGCCCTGAGGCATGCCCTTTCTCCA 1368  
|||||  
Db 2452 GTCTAGCTGGCTATCTCGCCCTTACTAACCCGCCCTGAGGCATGCCCTTTCTCCA 2511  
QY 1369 GCACACAAGCATTGGGGCAGCTGGAAATATTGGTTCCAGGCTCCTGTCTCTGGACTT 1428  
|||||  
Db 2512 GCACACAAGCATTGGGGCAGCTGGAAATATTGGTTCCAGGCTCCTGTCTCTGGACTT 2571  
QY 1429 CAGATCTCGGGGAGCCCTCCCGCCCTGAAATCCCTGGCTTAGCTTACCTTCTGCTGT 1488  
Db 2572 CAGATCTCGGGGAGCCCTCCCGCCCTGAAATCCCTGGCTTAGCTTACCTTCTGCTGT 2631  
QY 1489 GCACCTAAAAACCTCAGGTCAGACTAGGAAAGAGTTTGTATTTTGTGAATG 1548  
|||||  
Db 2632 GCACCTAAAAACCTCAGGTCAGACTAGGAAAGAGTTTGTATTTTGTGAATG 2691  
QY 1549 GAGTCTCGTCTGTGCGCCAGGCTGAGGTGAGTGTGCAATCTCCGCTCACTACAACCT 1608  
|||||  
Db 2692 GAGTCTCGTCTGTGCGCCAGGCTGAGGTGAGTGTGCAATCTCCGCTCACTACAACCT 2751  
QY 1609 CCACCTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCCGGAGTAGCTGGGACTATAG 1668  
Db 2752 CCACCTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCCGGAGTAGCTGGGACTATAG 2811  
QY 1669 GTGTGTACCATCACACCTGGCTAAATTTTGTATTTTGTAGACACAGGGTTTCGCCATG 1728  
|||||  
Db 2812 GTGTGTACCATCACACCTGGCTAAATTTTGTATTTTGTAGACACAGGGTTTCGCCATG 2871

QY 1729 TTGCCCAAGCTGCTTGAATTCCTGAGCTCAAGCAACCTGCGGCTCGGCCTCCCAA 1788  
|||||  
Db 2872 TTGCCCAAGCTGCTTGAATTCCTGAGCTCAAGCAACCTGCGGCTCGGCCTCCCAA 2931  
QY 1789 GTACTGGGATTTACAGCAGAAAGCCATGCCAGGCTAGATGTCTTATCCCAATCCT 1848  
|||||  
Db 2932 GTACTGGGATTTACAGCAGAAAGCCATGCCAGGCTAGATGTCTTATCCCAATCCT 2991  
QY 1849 TTGGCAGCATCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCT 1908  
|||||  
Db 2992 TTGGCAGCATCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCT 3051  
QY 1909 GGTTTAAAGCCAGATPAAGAGAAATCCCTTTCTAGGTTTGGATGTGTGTG 1962  
|||||  
Db 3052 GGTTTAAAGCCAGATPAAGAGAAATCCCTTTCTAGGTTTGGATGTGTGTG 3105

RESULT 4  
AX236082 1116 bp DNA linear PAT 26-SEP-2001  
LOCUS Sequence 13 from Patent WO0164887.  
DEFINITION AX236082  
ACCESSION AX236082  
VERSION AX236082.1 GI:15795889  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1116)  
AUTHORS Meyers, R.A.  
TITLE Patent: WO 0164887-A 13 07-SEP-2001;  
JOURNAL Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source  
1..1116  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
124..699  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC88255.1"  
/db\_xref="GI:15795890"  
translation="MLLGDYGVKTCFLIQKDGAFLSGTFIATYIDIRPKVTVVDG  
VRVKLIWDTAGQERFSVTHAYRDAQLLLYDITNKSFDNIRAMLTETHEYAQR  
DVIINLGNKADMSSEVRIRSEDGETLAREYGVPELTSAKTGMNVELAFLAIAKELK  
YRACHQADEPSFIIRDYVESQKRSQCSFM"  
BASE COUNT 250 a 313 c 322 g 221 t  
ORIGIN

Query Match 40.6%; Score 1086; DB 6; Length 1116;  
Best Local Similarity 99.5%; Pred. No. 5.1e-280;  
Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 37 GGGACATAGCGGACGCCAGCGCGTGGCCACCCGGATGCGAGGCCCGCCAGCGCT 96  
Db 23 CGGTCCGAGCGGACGCCAGCGCGTGGCCACCCGGATGCGAGGCCCGCCAGCGCT 82  
QY 97 CCCCAGCTTGCAGTCCGAGCTCAGGCGCAAGGTGATGCTTCTGGGAGACACAG 156  
Db 83 CCCCAGCTTGCAGTCCGAGCTCAGGCTCAGGCGCAAGGTGATGCTTCTGGGAGACACAG 142  
QY 157 GGTTCGGCAAAACATGTTTCTTCATCAATTCAAAGACGGGGCTTCTCTCCGGAACCT 216  
Db 143 GGTTCGGCAAAACATGTTTCTTCATCAATTCAAAGACGGGGCTTCTCTCCGGAACCT 202  
QY 217 TCATAGCCACCGTCCGCATAGACTTCCAGGAACAAGGTGGTACTGTGGATGGCGTAGAG 276  
Db 203 TCATAGCCACCGTCCGCATAGACTTCCAGGAACAAGGTGGTACTGTGGATGGCGTAGAG 262  
QY 277 TGAAGCTGCAGATCTGGGACACCGCTGGGAGGAAGGTTCCGAAGCGTCAACCCATGCTT 336  
Db 263 TGAAGCTGCAGATCTGGGACACCGCTGGGAGGAAGGTTCCGAAGCGTCAACCCATGCTT 322

QY 337 ATTACAGATGCTCAGGCGCTTCTTCTGCTATGACATCACCACAAATCTTCTTCG 396  
|||||  
Db 323 ATTACAGATGCTCAGGCGCTTCTTCTGCTATGACATCACCACAAATCTTCTTCG 382  
|||||  
QY 397 ACAACATCAGGCGCTTCTCCTCAGATTCATGAGTATGCCAGAGGAGCTGGTGATCA 456  
|||||  
Db 383 ACAACATCAGGCGCTTCTCCTCAGATTCATGAGTATGCCAGAGGAGCTGGTGATCA 442  
|||||  
QY 457 TGCTGTAGGCAACAAGCGGATATGAGCAGCGAAAGAGTGATCCGPTCCGAGAGCGGAG 516  
|||||  
Db 443 TGCTGTAGGCAACAAGCGGATATGAGCAGCGAAAGAGTGATCCGPTCCGAGAGCGGAG 502  
|||||  
QY 517 AGACCTTGGCCAGGAGTACGGTTCCTTCTTCTGAGACAGCGCCAAAGCTGGCATGA 576  
|||||  
Db 503 AGACCTTGGCCAGGAGTACGGTTCCTTCTTCTGAGACAGCGCCAAAGCTGGCATGA 562  
|||||  
QY 577 ATGTGGAGTTAGCTTCTTCTGCCATCCCAAGGAAGTGAATACCGGGCGGGCATCAGG 636  
|||||  
Db 563 ATGTGGAGTTAGCTTCTTCTGCCATCCCAAGGAAGTGAATACCGGGCGGGCATCAGG 622  
|||||  
QY 637 CGGATGAGCCAGCTTCCAGATCCGAGATATGTAGAGTCCCAAGAGCGGTCCAGCT 696  
|||||  
Db 623 CGGATGAGCCAGCTTCCAGATCCGAGATATGTAGAGTCCCAAGAGCGGTCCAGCT 682  
|||||  
QY 697 GCTGCTCTTCTATGTAATCCAGGGGCGAGAGGAGGCTCTGGAGGCACACAGATGC 756  
|||||  
Db 683 GCTGCTCTTCTATGTAATCCAGGGGCGAGAGGAGGCTCTGGAGGCACACAGATGC 742  
|||||  
QY 757 AGCTTCTCCCTCCAGGCTTCTTATCCAGAGGCTGAGCAATGGGGGAGAAAGATG 816  
|||||  
Db 743 AGCTTCTCCCTCCAGGCTTCTTATCCAGAGGCTGAGCAATGGGGGAGAAAGATG 802  
|||||  
QY 817 GAGGACTCACTGCACAGCCGCTTCTTAGCAGGAGCTATCTCCAATCTTACTTGAGTT 876  
|||||  
Db 803 GAGGACTCACTGCACAGCCGCTTCTTAGCAGGAGCTATCTCCAATCTTACTTGAGTT 862  
|||||  
QY 877 CTTGCGCTTCCCGCATCCACAGGAGGTAAACACTTAGCTTTATTTTAACTAGTAC 936  
|||||  
Db 863 CTTGCGCTTCCCGCATCCACAGGAGGTAAACACTTAGCTTTATTTTAACTAGTAC 922  
|||||  
QY 937 ATAATTTAATACCAAAAAGCGCTTGATCCCAAAAACCGAGGCTGGGAGCTAGTGG 996  
|||||  
Db 923 ATAATTTAATACCAAAAAGCGCTTGATCCCAAAAACCGAGGCTGGGAGCTAGTGG 982  
|||||  
QY 997 CCCTTTTGTCTTAGACACTTGGGGGCGCGCCCTCCTCTTAAGCATACAAAGTGCT 1056  
|||||  
Db 983 CCCTTTTGTCTTAGACACTTGGGGGCGCGCCCTCCTCTTAAGCATACAAAGTGCT 1042  
|||||  
QY 1057 GTTGCTCCAGCTCAGCCCAAGGGGACACAGATGCACCTTTGGGGGTGAGGCGAGTATGA 1116  
|||||  
Db 1043 GTTGCTCCAGCTCAGCCCAAGGGGACACAGATGCACCTTTGGGGGTGAGGCGAGTATGA 1102  
|||||  
QY 1117 CTCATCGCACCTT 1130  
|||||  
Db 1103 CTCATCGCACCTT 1116  
|||||  
RESULT 5  
BC016615 1700 bp mRNA linear PRI 05-NOV-2001  
LOCUS  
DEFINITION Homo sapiens, similar to RAB37, member of RAS oncogene family,  
clone MGC:21391 IMAGE:4520191, mRNA, complete cds.  
ACCESSION BC016615  
VERSION BC016615.1 GI:16741620  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1700)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (31-OCT-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: DCTD/DTT  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalona@bcm.tmc.edu](mailto:villalona@bcm.tmc.edu)  
Villalón, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 28 Row: j Column: 15  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.  
Location/Qualifiers  
1. .1700  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:21391 IMAGE:4520191"  
/tissue\_type="Prostate, adenocarcinoma."  
/clone\_lib="NIH\_MGC\_91"  
/lab\_host="DH10B"  
/notes="Vector: pCMV-SPORT6"  
3. .674  
/codon\_start=1  
/product="Similar to RAB37, member of RAS oncogene family"  
/protein\_id="AAH16615.1"  
/db\_xref="GI:16741621"  
/translation="MTGTCGAVATRDGEAPSPSPSYDLTKYMLLGDGTGVGKTC  
FTIQKDGAFSLGTFATIDFNRKVVTVGVKVLQIWDTAGQRFRRSVTHAYVRD  
AQAALLLYDITNKSPDNIRAWLTIHEYAQRDVIIMLLGNKADMSRVRSEDGET  
LAREYGVPELTSKATGMVVELAFLAIKELKYRAGHQADPEFQIRDYVESQKRKSS  
CCSPW"

FEATURES  
Source  
1. .1700  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:21391 IMAGE:4520191"  
/tissue\_type="Prostate, adenocarcinoma."  
/clone\_lib="NIH\_MGC\_91"  
/lab\_host="DH10B"  
/notes="Vector: pCMV-SPORT6"  
3. .674  
/codon\_start=1  
/product="Similar to RAB37, member of RAS oncogene family"  
/protein\_id="AAH16615.1"  
/db\_xref="GI:16741621"  
/translation="MTGTCGAVATRDGEAPSPSPSYDLTKYMLLGDGTGVGKTC  
FTIQKDGAFSLGTFATIDFNRKVVTVGVKVLQIWDTAGQRFRRSVTHAYVRD  
AQAALLLYDITNKSPDNIRAWLTIHEYAQRDVIIMLLGNKADMSRVRSEDGET  
LAREYGVPELTSKATGMVVELAFLAIKELKYRAGHQADPEFQIRDYVESQKRKSS  
CCSPW"

BASE COUNT 437 a 444 c 460 g 359 t  
ORIGIN  
Query Match 35.1%; Score 938.4; DB 9; Length 1700;  
Best Local Similarity 99.9%; Pred. No. 2.2e-240;  
Matches 939; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1735 AGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTCGCGGCTCGGCTCCCAAGTACTG 1794  
|||||  
Db 752 AGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTCGCGGCTCGGCTCCCAAGTACTG 811  
|||||  
QY 1795 GGATTACACGACAGGACCATGCCAGCTAGATGTGTATCCCAATCCTTTGGCA 1854  
|||||  
Db 812 GGATTACACGACAGGACCATGCCAGCTAGATGTGTATCCCAATCCTTTGGCA 871  
|||||  
QY 1855 GGATGCGAGCTCCACAGCGGATTTCTTCAAGCAGCTCAAGTGTTTAGCCCTCCTGGGTTA 1914  
|||||  
Db 872 GGATGCGAGCTCCACAGCGGATTTCTTCAAGCAGCTCAAGTGTTTAGCCCTCCTGGGTTA 931  
|||||  
QY 1915 AGAGCCAGATAAGGAGAAATCCCTTTCTAGTGTGGAAATGTGTGTGAAAAAAGAGA 1974  
|||||  
Db 932 AGAGCCAGATAAGGAGAAATCCCTTTCTAGTGTGGAAATGTGTGTGAAAAAAGAGA 991  
|||||  
QY 1975 AATCCCTGGCTCTCGAGCTGGTGGGAGACAAGTAAGCAAAACCTCCCTGACATGTAT 2034  
|||||  
Db 992 AATCCCTGGCTCTCGAGCTGGTGGGAGACAAGTAAGCAAAACCTCCCTGACATGTAT 1051  
|||||  
QY 2035 CCCTTTGACCCCAAGCTCTGCCCTCCCTGACACCATGCCCTTTCCCTTTAACTTCTC 2094  
|||||  
Db 1052 CCCTTTGACCCCAAGCTCTGCCCTCCCTGACACCATGCCCTTTCCCTTTAACTTCTC 1111  
|||||

```
QY 2095 AAACAGATACCGGCGCTAAACTGCTTTTACCTCCCTCCCTACTGAGTCAGGTTAGTGCT 2154
Db 1112 AAACAGATACCGGCGCTAAACTGCTTTTACCTCCCTCCCTACTGAGTCAGGTTAGTGCT 1171
QY 2155 GGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAGTCATGTGGG 2214
Db 1172 GGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAGTCATGTGGG 1231
QY 2215 TATGCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTCCTTGGTTCACATATCTCC 2274
Db 1232 TATGCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTCCTTGGTTCACATATCTCC 1291
QY 2275 CAAAGCTCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTTCATGACCTGCAGGC 2334
Db 1292 CAAAGCTCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGGTTCATGACCTGCAGGC 1351
QY 2335 ATCTTTACTGACGCTCTGCGGCCCTGGAGGGGAGAGGGGAGGAAGTATGCGCTGC 2394
Db 1352 ATCTTTACTGACGCTCTGCGGCCCTGGAGGGGAGAGGGGAGGAAGTATGCGCTGC 1411
QY 2395 ACATTTCTGAGGCTACTGCATTTGCTTTCAAGSCAGAAATCTTGCTCTGAGCAGTCAGCG 2454
Db 1412 ACATTTCTGAGGCTACTGCATTTGCTTTCAAGSCAGAAATCTTGCTCTGAGCAGTCAGCG 1471
QY 2455 GCTCCAGTTTGGGCGCGATAAGGAAGTTCTCGTGCGCTCCCTCAGGCAGAGCAGGGAGG 2514
Db 1472 GCTCCAGTTTGGGCGCGATAAGGAAGTTCTCGTGCGCTCCCTCAGGCAGAGCAGGGAGG 1531
QY 2515 AGGCTGACATTCAGCTCTCTTCTGGGCGCCAAAGCAGGTTGACAGAGATCCAAATCCCAT 2574
Db 1532 AGGCTGACATTCAGCTCTCTTCTGGGCGCCAAAGCAGGTTGACAGAGATCCAAATCCCAT 1591
QY 2575 AGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAATTAACATTCACATTTTGGAAA 2634
Db 1592 AGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAATTAACATTCACATTTTGGAAA 1651
QY 2635 GCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
Db 1652 GCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1691
RESULT 6
LOCUS AX236084 AX236084 576 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 15 from Patent WO0164887.
ACCESSION AX236084
VERSION AX236084.1 GI:15795891
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Meyers,R.A.
TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins
JOURNAL Patent: WO 0164887-A 15 07-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source
1..576
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 141 a 148 c 167 g 120 t
ORIGIN
Query Match 21.5%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.7e-143;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 138 ATGCTCTCTGGGAGACACAGCGCTCGCAAAACATGTTTCTGATCCATTCAAAGACGGG 197
Db 1 ATGCTCTCTGGGAGACACAGCGCTCGCAAAACATGTTTCTGATCCATTCAAAGACGGG 60
```

```
QY 198 GCCTTCTCTCGGAACCTTTCATAGCCACCGCTGGCATAGACTTCAGGAACAGGTGGTG 257
Db 61 GCCTTCTCTCGGAACCTTTCATAGCCACCGCTGGCATAGACTTCAGGAACAGGTGGTG 120
QY 258 ACTGTGGATGGCGTGAAGCTGACAGTCTGGGACACCGCTGGGAGGAAACGGTTC 317
Db 121 ACTGTGGATGGCGTGAAGCTGACAGTCTGGGACACCGCTGGGAGGAAACGGTTC 180
QY 318 CCAAGCGTCACCCATGCTTTATTTACAGAGATGCTCAGCGCTTGGTCTGCTGTATGACATC 377
Db 181 CCAAGCGTCACCCATGCTTTATTTACAGAGATGCTCAGCGCTTGGTCTGCTGTATGACATC 240
QY 378 ACCAACAATCTTCTTTCGACAACATCAGGCGCTGGCTCACCTGAGATTCATGATATGCC 437
Db 241 ACCAACAATCTTCTTTCGACAACATCAGGCGCTGGCTCACCTGAGATTCATGATATGCC 300
QY 438 CAGAGGAGCTGGTGATCATGCTGTAGGCAACAAGCGCGATATGAGCAGCGAAAGAGTG 497
Db 301 CAGAGGAGCTGGTGATCATGCTGTAGGCAACAAGCGCGATATGAGCAGCGAAAGAGTG 360
QY 498 ATCCGTTCCGAAGACGAGAGACCTTGGCCACGAGGAGTACGGTCTCCCTTCCCTGGAGACC 557
Db 361 ATCCGTTCCGAAGACGAGAGACCTTGGCCACGAGGAGTACGGTCTCCCTTCCCTGGAGACC 420
QY 558 AGCGCCAAGACGTCGATGAATGTGGAGTTAGCTTTTCTGGCCATCGCCAGGAACCTGAAA 617
Db 421 AGCGCCAAGACGTCGATGAATGTGGAGTTAGCTTTTCTGGCCATCGCCAGGAACCTGAAA 480
QY 618 TACCGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
Db 481 TACCGGGCGGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540
QY 678 CAGAAGAACGCTCCAGCTGCTGCTCTTCATGTGA 713
Db 541 CAGAAGAACGCTCCAGCTGCTGCTCTTCATGTGA 576
RESULT 7
LOCUS AF233582 AF233582 690 bp mRNA linear ROD 02-MAY-2000
DEFINITION Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
ACCESSION AF233582
VERSION AF233582.1 GI:7677421
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 690)
AUTHORS Yu,S., Bennett,M.K., Luo,Y., Young,C., Shen,M., Rossi,A.B., Huang,B.C.,
Masuda,E.S., Luo,Y., Payan,D.G. and Scheller,R.H.
TITLE Rab37 is a novel mast cell specific GTPase localized to secretory
granules
JOURNAL FEBS Lett. 470 (1), 61-64 (2000)
MEDLINE 20189834
PUBMED 10728846
REFERENCE 2 (bases 1 to 690)
AUTHORS Luo,Y., Huang,B.C.B., Yu,S., Shen,M. and Masuda,E.S.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
Avenue, South San Francisco, CA 94080, USA
FEATURES
Location/Qualifiers
source
1..690
/organism="Mus musculus"
/db_xref="taxon:10090"
gene
1..690
/feature="Rab37"
CDS
1..672
/feature="Rab37"
/codon_start=1
/product="GTPase Rab37"
/protein_id="AAF67162.1"
/db_xref="GI:7677422"
```









```
CDS
/gene="v46133"
4..576
/gene="v46133"
/note="RAS-related protein"
/codon_start=1
/product="RAB-26"
/protein_id="BA84707.1"
/db_xref="GI:5931612"
/translation="MLVGSYGKTLVRFKDGAFAGTFFSTVIGIDPRNKVLVDVG
VKVQLMDTAGQERFSTHAYYRDAHALLLYDTNKASFDNQAWLTHEYAQH
DVALMLGNKVSASHERVVKREDGEKLAKEYGLPFWEISAKTGLNVDLAFTAIKELK
QRSMKAPSEPRFLHDYVYKREGASCCRP"
BASE COUNT 273 a 379 c 395 g 273 t
ORIGIN
Query Match 11.4%; Score 304.6; DB 9; Length 1320;
Best Local Similarity 71.1%; Pred. No. 2e-70;
Matches 403; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
QY 135 GTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATCAAGAC 194
D 1 GTCATGCTGTGGGACTCGGGTGTGGGAAGACCTCTGCTGGTGGCATTCAGGAT 60
QY 195 GGGCCCTTCTGTCGGAACCTTCATAGCACCGTCGGGATAGACTTCAGGAACAAGGTG 254
D 61 GGTGCTTCTGGGGGAGCTTCATCTCCACCGTAGGCATTGACTTCCGGAAACAAGTT 120
QY 255 GTGACTGTGGATGCGTGAAGCTGAAGCTGTGGACACCGCTGGGACGAACGG 314
D 121 CTGACGTGGATGTGGAAGTGAAGCTGCATGTGGACACAGCTGTGAGGACGG 180
QY 315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTTGTCTGTGATGAC 374
D 181 TTCCGAGTGTACCCATGCTACTACCGGATGCTCATGCTGTGCTGTCTACGAT 240
QY 375 ATCAACAATAATCTTTTCGACACATCAGGCGCTGGCTACTGAGATTCATGATAT 434
D 241 GTCACCAACAAGGCTCTCTTTGACAACTCAGGCGCTGGCTGACCGAGATCCACGATAC 300
QY 435 GCCCAGAGGACGTGTGTGATCTGCTAGGCAACAAGCGGATATGAGCAGCAAGA 494
D 301 GCCCAGACAGCTGTGGGCTCATGCTGTGGGACAAAGTGGACTTGCCCATGACGT 360
QY 495 GTGATCCGTTCCGAAGACGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCCTGGAG 554
D 361 GTGTGAAGAGGAGGACGGGAGAGCTGGCCAAAGGATGATGACTGCCCTTCATGGAG 420
QY 555 ACCAGCGCAAGACTGGCATGAATGTGGAGTTAGCCTTCTTGGCCATCGCAAGGAAGT 614
D 421 ACCAGCGCAAGACGGGCTCAAGTGGACTTGGCTTTCACAGCCATAGCAAGAGTGT 480
QY 615 AAATACCGGGCGGCGCATCAGCGCGGATGAGCCCAAGCTTCCAGATCCGAGACTATGAGAG 674
D 481 AAGCAGCGCTCCATGAAGGCTCCACGAGGACGGCGCTTCCGGCTGCATGATTAGTTAAG 540
QY 675 TCCAGAGAAGAGCGCTCCAGCTGTGC 701
D 541 AGGAGGCTCAGGGGCGCTCCTGTGC 567

RESULT 11
AF498952
LOCUS
DEFINITION Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete cds.
ACCESSION AF498952
VERSION AF498952.1
KEYWORDS GI:20379079
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
```

AUTHORS Puhl, H.L. III, Ikeda, S.R. and Aronstam, R.S.  
TITLE Homo sapiens RAB family small GTP binding protein RAB26  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 573)  
AUTHORS Puhl, H.L. III, Ikeda, S.R. and Aronstam, R.S.  
TITLE Direct Submission  
JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research Institute, One Guthrie Square, Sayre, PA 18840, USA

FEATURES  
Location/Qualifiers  
source  
1..573  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
1..573  
/gene="RAB26"  
1..573  
/gene="RAB26"  
/note="RAB family member"  
/codon\_start=1  
/product="small GTP binding protein RAB26"  
/protein\_id="AA021100.1"  
/db\_xref="GI:20379080"  
/translation="MLVGSYGKTLVRFKDGAFAGTFFSTVIGIDPRNKVLVDVG  
VKVQLMDTAGQERFSTHAYYRDAHALLLYDTNKASFDNQAWLTHEYAQH  
DVALMLGNKVSASHERVVKREDGEKLAKEYGLPFWEISAKTGLNVDLAFTAIKELK  
QRSMKAPSEPRFLHDYVYKREGASCCRP"

BASE COUNT 117 a 151 c 191 g 114 t  
ORIGIN

Query Match 11.3%; Score 303.2; DB 9; Length 573;  
Best Local Similarity 71.1%; Pred. No. 4e-70;  
Matches 401; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 138 ATGCTTCTGGGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATCAAGACGG 197  
D 1 ATGCTGTGGGACCTCGGGTGTGGGAAGACCTGTCTGTGCTGCTCAAGGATG 60  
QY 198 GCCTTCTGTCGGGAACCTTCATAGCACCGTCGGCATAGACTTCAGGAACAAGGTGGT 257  
D 61 GCTTCTGTCGGGGACCTTCATCTCCACCGTAGGCATTGACTTCCGGAACAAGTTCTG 120  
QY 258 ACTGTGGATGCGTGTGAGAGTGAAGCTGTGGACACCGCTGGGACGAACGGTTC 317  
D 121 GAGCTGATGCTGTGAAGTGAAGCTGCATGTGGACACAGCTGTGAGGACGGTTC 180  
QY 318 CGAAGCCTCACCCTGCTTATTACAGAGATGCTCAGGCGCTTGTCTGTGATGATC 377  
D 181 CGCAGTGTACCCATGCTTACTACCGGATGCTCATGCTGTGCTGTCTACGATGTC 240  
QY 378 ACCAACAATCTTCTTCGACACATCAGGCGCTTGGCTCACTGAGATTCATGATGTC 437  
D 241 ACCAACAAGCGCTTCTTGAACAATCCAGGCGCTGGCTGACCGAGATCCACGAGTAGCC 300  
QY 438 CAGAGGACCTGTGTGATCATGCTGTAGGCAACAAGCGGATATGAGCAGCAAGAGTG 497  
D 301 CAGCAGACCTGGGCTCATGCTGTGGGAACAAGTGGACTTGCCCATGAGCGTGTG 360  
QY 498 ATCCGTTCCGAAGACGAGAGACCTTGGCCAGGAGTACGGTGTTCCTTCCTGGAGACC 557  
D 361 GTGAAGAGGAGGAGCGGGGAGAGCTGGCCAAAGGAGTATGGACTTGCCTTCATGAGAAC 420  
QY 558 AGGCCAAGACTGGCATGATGTGGATGTAGCTTTCCTGCCCACCGCAAGACTCAAA 617  
D 421 AGGCCAAGACGGGCGCTCAACGTTGGACTTGCCTTTCACAGCCATAGCAAGAGTGAAG 480  
QY 618 TACCGGCGGGGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677  
D 481 CAGGCTCCATGAAGGCTCCAGGAGCGCGCTTCCGGCTGCATGATGATGATAGAGG 540  
QY 678 CAGAAGAAGCGCTCCAGCTGCTGC 701  
D 541 GAGGTCGAGGGGCGCTCCTGCTGC 564

RESULT 12	AR062279	Sequence 2 from patent US 5843717.	1340 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	AR062279					
DEFINITION	Sequence 2 from patent US 5843717.					
ACCESSION	AR062279					
VERSION	AR062279.1	GI:5989970				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1340)					
AUTHORS	Hillman,J.L. and Guegler,K.J.					
TITLE	Rab protein					
JOURNAL	Patent: US 5843717-A 2 01-DEC-1998;					
FEATURES	Location/Qualifiers					
Source	1..1340					
BASE COUNT	276 a 387 c 403 g 274 t					
ORIGIN	/organism="unknown"					
	Query Match 10.9%; Score 292.6; DB 6; Length 1340;					
	Best Local Similarity 70.9%; Pred. No. 3.4e-67;					
	Matches 416; Conservative 0; Mismatches 169; Indels 2; Gaps 2;					
QY	116	CTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT	175			
Db	21	CTACGACGTGCGCTTCAAGTCTATGCTGGTGGGGACTCGGNGTGGGAAGACCTGTCT	80			
QY	176	CTGGA-TCCAAATCAAGACGGGGCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCA	234			
Db	81	GC7GGTGGCATTCAGGATGGTGTCTTCTGTGGGGGACCTTCATCTCCACCGT-AGCA	139			
QY	235	TAGACTTCAGGAACAAGGTGGTCACTGTGGATGCGTGGAGAGTGAAGCTGCAGATCTGG	294			
Db	140	TTGACTTCCGGAACAAGTCTCGACGTGGATGGTGTGGAAGGTGAAGCTGCAGATGTGG	199			
QY	295	ACACCGCTGGGACGACGCTTCCGAAGCGTCAACCATGCTTATTACAGAGATGCTCAGG	354			
Db	200	ACACAGCTGTCAGGACGCTTCCGACGTGTTACCCATGCTACTACCGGATGCTCATG	259			
QY	355	CCCTGCTTCTGCTGTATGATACATCACCAAAATCTTCTTCGACACATCAGGCGCTGC	414			
Db	260	CTCTGTGCTGCTCTACGATGTCCACCAACAGGCGCTCTTGTGACACATCAGGCGCTGC	319			
QY	415	TCACTGAGATTCATGATGATGTCGACAGGAGCGTGTGATCATGCTAGGCAACAAGG	474			
Db	320	TGACCGAGATCCAGAGTACGCCACGACGCGTGGCGCTCATGCTGGGGAACAAGG	379			
QY	475	CGGATATGACGACGAAAGAGTATCCGTTCCGAAGACGAGAGACCTTGGCCAGGGAGT	534			
Db	380	TGGACTCTGCCCATGAGCGGTGTGTTGAAGAGGAGGAGCGGGAGAAAGCTGGCCAAAGG	439			
QY	535	ACGCTGTTCCCTTCTTGGAGACACGCGCCAGACTGCGCATGAATGTGGAGTTAGCCTTC	594			
Db	440	ATGAGCTGCGCTTCATGAGACACGCGCCCAAGACGCGCTTCAACGTGGACTTGGCTTCA	499			
QY	595	TGCCCATCGCAAGGAACCTGAATACCGGCGCGGATCAGCGGATCAGCCAGCTTCC	654			
Db	500	CAGCCATAGCAAGGAGTGTGAACACGCGCTCCATGAAGGCTCCACGAGCGCGCTTCC	559			
QY	655	AGATCCGAGACTATGTAGATGCTCCAGGAAGAAGCGCTCCAGCTGCTGC 701				
Db	560	GGCTGATGATTAAGTTAAGAGGAGGCTGAGGGCGCTTCTCTGCTGC 606				
RESULT 13	AY061826	Sequence 2536 bp	2536 bp	mRNA	linear	INV 10-NOV-2001
LOCUS	AY061826					
DEFINITION	Drosophila melanogaster GH21984 full length cDNA.					
ACCESSION	AY061826					
VERSION	AY061826.1	GI:16902019				
KEYWORDS	FLI_CDNA.					

SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 2536)
	Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J., Champre,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J., Nuncio,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
TITLE	Direct Submission
JOURNAL	Submitted (02-NOV-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
COMMENT	Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory Berkeley, CA 94720
	This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (http://fruitfly.berkeley.edu) or send email to cdna@fruitfly.berkeley.edu.
FEATURES	Location/Qualifiers
source	1..2536
	/organism="Drosophila melanogaster"
	/strain="y; cn bw sp"
	/db_xref="taxon:7227"
	/map="78C9-78D1"
gene	1..2536
	/gene="CG7605"
	/note="alignment with genomic scaffold AE003594"
	/db_xref="FLYBASE:FBgn0037072"
	1172..2338
CDS	/gene="CG7605"
	/note="Longest ORF"
	/codon_start=1
	/product="GH21984p"
	/protein_id="AAL27637.1"
	/db_xref="GI:16902020"
	/db_xref="FLYBASE:FBgn0037072"
	/translation="MASTAVGLGGEGDPGAGPPAGSAHPDDASSDDVDFDAETQARIELRRRPFPGSGSYNPPAAPASVSASITTTTQQOQHNPSSHHQSHQPSHHHHHHSQSLTSGSHYHDDAIMAPVRSATGYPGRPSREAMQYATGDTDDDDTGDWRSYRYDEVDMHPAPSNHQPFDTVNNKTIILGDSGVGKTSFLVNTGFEKLG SFSATGIALTKNVVVDGTRVKLIQIWDTAGOERFSRTHAYRDHALLLLDVNTK TYDNIIRAWLGEIRAEQDVIVLIGNKADCSGRQVRDEGRGRGRHNVPFMET SAKTGLNVELSFATAVQLKSRGYEHDDGDFRNVHDFVRNTRKARSVCAACRN"
BASE COUNT	831 a 621 c 624 g 460 t
ORIGIN	
	Query Match 8.8%; Score 234.2; DB 3; Length 2536;
	Best Local Similarity 65.7%; Pred. No. 1.8e-51;
	Matches 357; Conservative 0; Mismatches 182; Indels 3; Gaps 1;
QY	131
Db	1744
QY	191
Db	1804

SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	Stapleton,M., Bravo,C., Dorsett,V., Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M. and Celnikier,S.
TITLE	Direct Submission
JOURNAL	Submitted (02-NOV-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
COMMENT	Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory Berkeley, CA 94720 This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (http://fruitfly.berkeley.edu) or send email to cdna@fruitfly.berkeley.edu.
FEATURES	Location/Qualifiers
source	1..2536
gene	/organism="Drosophila melanogaster" /strain="y; cn bw sp" /db_xref="taxon:7227" /map="78C9-78D1" 1..2536
CDS	/gene="CG7605" /note="alignment with genomic scaffold AE003594" /db_xref="FLYBASE:FBgn0037072" 1172..2338 /note="longest ORF" /product="GH21984p" /protein_id="AAL27637.1" /db_xref="GI:16902020" /db_xref="FLYBASE:FBgn0037072" /translation="MASTAVLGEGGDFGAGGPPAGSHPDASMSDDVDFDAETT QARIEELRRPFDDGSYNPAPASVSASITTTQOOOHHNPSSHHHSHHSHH HHHHHSLSLTGSHHYHDDAIPQVRSATGYPGRPSREAMQYAYGTDDYDYN DQWRSYRDEVDMPAPNAHQPFDDTVNHKTILGDSGVKTSFLVRYNTEGFRLG SFSATVGIALTNKVVVDGTRVKIOWTAGOERFSVTHAYVRDAHALLLVDVTK TYYDNIRALGELREYAEQDVILVILGNKADCSGSEKQVKEGRLGHEHNPVEMET SAKTGLNVELSFTAVARQLKSGRYEHGDDGRFVHDFVNDTKARVCAQCRNM"
BASE COUNT	831 a 621 c 624 g 460 t
ORIGIN	
Query Match 8.8%; Score 234.2; DB 3; Length 2536; Best Local Similarity 65.7%; Pred. No. 1.8e-51; Matches 357; Conservative 0; Mismatches 183; Indels 3; Gaps 1;	
QY	131 CAAGGTGATGCTTCTGGGAGACACAGGCTCGGCAAAACATGTTCTGATCCAAATCAA 190
Db	1744 CAAGACCATCTCTCTGGGATTCGGCGCTGGGAAAGACCTCTCTCTGTCATACAA 1803
QY	191 AGACGGGCGCTTCTCTGCGGAACCTTCATAGCCACCGTCGCGATAGACTTCAGGAACAA 250
Db	1804 TACGGGCGAGTTCGAGCTGGGTTCTCTTTCCGCCACACATGGGATTCGCGTTACGAACAA 1863



```
polyA_signal complement(10783..10788)
polyA_site /gene="FBLN1"
polyA_site /gene="FBLN1"
repeat_region 11114..11149
repeat_region /note="MIR repeat: matches 109..144 of consensus"
repeat_region 11527..11572
repeat_region /note="MIR repeat: matches 75..120 of consensus"
repeat_region 11651..11934
repeat_region /note="L2 repeat: matches 1992..2328 of consensus"
repeat_region 12001..12270
repeat_region /note="L2 repeat: matches 2454..2750 of consensus"
misc_feature 12337..13158
misc_feature /gene="FBLN1"
misc_feature /note="match: STS: Em:X53742"
misc_feature <12345..13158
misc_feature /gene="FBLN1"
misc_feature /product="bk941f9.2 (Fibulin 1 (isoform B))"
misc_feature /note="match: CDNAS: Em:AB007943 Em:AF078528 Em:Y14610
misc_feature Em:AF011357 Em:X53742 Em:AF019887; match: STS: Em:R78704
misc_feature Em:R33281 Em:R32733 Em:H01971 Em:H02060"
misc_feature /evidence=not_experimental
misc_feature <12345..12449
misc_feature /gene="FBLN1"
misc_feature /note="match: proteins: Sw:P23143"
misc_feature /codon_start=1
misc_feature /evidence=not_experimental
misc_feature /product="bk941f9.2 (Fibulin 1 (isoform B))"
misc_feature /protein_id="CAB62961.1"
misc_feature /db_xref="GI:6572284"
misc_feature /translation="KSKKGRQNTPAGSSKEDCRVLPWKQGLETHLDA"
misc_feature 12449..12703
misc_feature /gene="FBLN1"
misc_feature /note="match: STS: Em:G59664"
misc_feature 12450..12849
misc_feature /gene="FBLN1"
misc_feature /note="match: STS: Em:G11168"
misc_feature complement(12733..13139)
misc_feature /note="match: STS: Em:H01971"
misc_feature 12895..13143
misc_feature /note="L2 repeat: matches 2419..2705 of consensus"
misc_feature complement(12974..13280)
misc_feature /note="match: STS: Em:G08093; match: STS: Em:G08093"
polyA_signal 13138..13143
polyA_signal /gene="FBLN1"
polyA_site 13158
polyA_site /gene="FBLN1"
polyA_site 13237..13536
polyA_site /note="AluSp repeat: matches 1..300 of consensus"
polyA_site 13974..14045
polyA_site /note="MIR repeat: matches 50..122 of consensus"
polyA_site 14352..14607
polyA_site /note="MIR repeat: matches 2..262 of consensus"
polyA_site 14791..15080
polyA_site /note="AluSx repeat: matches 1..290 of consensus"
polyA_site 15714..15956
polyA_site /note="MLTIC repeat: matches 300..451 of consensus"
polyA_site 16332..16625
polyA_site /note="MLTIC repeat: matches 1..310 of consensus"
polyA_site 16637..16815
polyA_site /note="MLTIB repeat: matches 1..190 of consensus"
polyA_site 16816..17125
polyA_site /note="AluSg repeat: matches 1..310 of consensus"
polyA_site 17126..17338
polyA_site /note="MLTIB repeat: matches 190..381 of consensus"
polyA_site 17506..17806
polyA_site /note="AluYb8 repeat: matches 1..318 of consensus"
polyA_site 17807..18030
polyA_site /note="7 copies 32 mer 64 conserved"
polyA_site 17808..18031
polyA_site /note="112 copies 2 mer aa 62 conserved"
polyA_site 18046..18303
```

```
repeat_region /note="HAL1 repeat: matches 351..597 of consensus"
repeat_region 18304..18597
repeat_region /note="AluSg repeat: matches 1..309 of consensus"
repeat_region 18598..18822
repeat_region /note="HAL1 repeat: matches 597..825 of consensus"
repeat_region 19070..19305
repeat_region /note="L1MC5 repeat: matches 6484..6730 of consensus"
repeat_region 19306..19621
repeat_region /note="AluSg repeat: matches 1..310 of consensus"
repeat_region 19622..20401
repeat_region /note="L1MC5 repeat: matches 6730..7471 of consensus"
repeat_region 20402..20711
repeat_region /note="AluSx repeat: matches 1..312 of consensus"
repeat_region 20712..20770
repeat_region /note="L1MC5 repeat: matches 7471..7531 of consensus"
repeat_region 20851..21158
repeat_region /note="AluSg repeat: matches 1..301 of consensus"

Query Match 8.0%; Score 214.8; DB 9; Length 127587;
Best Local Similarity 80.3%; Pred. No. 7.5e-46;
Matches 252; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1526 TTGTGTTTATTTTTTGAATGGAGTCGCTGTCTGCCCCAGGCTGAGGTGCAGTACT 1585
Db 59011 TTTGTTTTTCTTTTTTGAACGGAGTCTCACTGTCTCACCAGGCTGGAGTGCAATGGC 58952

QY 1586 GCAATCTCGCTCACTACAACCTCCACTCCCTGGGCTCAAGCATCCCTCCACCTCAGC 1645
Db 58951 GGCATCTGGCTTATGAGCCCTCTGCCCTCCGGGATCAAGTGATTCCTTCCTCAGC 58892

QY 1646 CCCCAGATAGCTGGGACTATAGGTGTGTTACCATCACACCTGGCTAATTTTTTGTATTTT 1705
Db 58891 CTCCCAAGTAGCTGGAATTACAGGTGCATACCACACACCTGGCTAATTTTTTATATTTT 58832

QY 1706 TGTACACACAGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTCAGCTCAGCAA 1765
Db 58831 AGTAGAGATGGGGTTTCAACCATGTTGCCAGGCTGGTCTTGAATTCCTCAGCTCAGCAA 58772

QY 1766 CTGCGCGGCTCGGCTCCCAAAGTACTGGGATTTACACGACGAAGCACCATGCCAGCGC 1825
Db 58771 TCTGCTGCTCGGCTCCCAAAGTCTCCGAGTATACAGCAGACCATGCCAGCGGCC 58712

QY 1826 TAGATGTGTTCTTAT 1839
Db 58711 TGGTTTTCTGTTTT 58698

RESULT 15
AC010352/c
LOCUS AC010352 102332 bp DNA linear PRI 18-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2022F20, complete sequence.
ACCESSION AC010352
VERSION AC010352.4 GI:7109394
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 102332)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102332)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 102332)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (29-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 102332)
```

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (18-APR-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Feb 29, 2000 this sequence version replaced gi:6606056.  
Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.4% of Sequence;  
Estimated Total Number of Errors is 1.5.  
STS Content:  
WI-12799 G24393  
SHGC-13554 G14788.

FEATURES  
source  
1. .102332  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2022F20"  
BASE COUNT 29578 a 22934 c 23227 g 26593 t  
ORIGIN

Query Match 8.0%; Score 214.2; DB 9; Length 102332;  
Best Local Similarity 82.6%; Pred. No. 1e-45;  
Matches 257; Conservative 0; Mismatches 53; Indels 1; Gaps 1;  
QY 1522 GAGTTTGTGTTTATTTTGAATGGAGTCGTTCTGTCGCCAGGCTGAGGTGCAG 1581  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
Db 22554 GATTTTATTTTATTTTGTGACACAGTCGCTCTGTCGCCAGGCTGGAATGCAG 22495  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
QY 1582 TAGTGAATCTCGGCTCACTACAACCTCCACTCCCTCGGGCTCAAGCGATCCTCCACCT 1641  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
Db 22494 TGGTGAATCTCGGCTCCCTGCAACCTCCACTTCCT -GGTTGAAGCGATTCTCTGCCT 22436  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
QY 1642 CAGCCCGGAGTAGTGGGACTATAGGTGTGTACCATCACACCTGGCTAAATTTTGTAT 1701  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
Db 22435 CAGCCTCCTGAGTAGTGGGACTACAGGCGTGTGCCACACACCCGGCTAAATTTTGTAT 22376  
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
QY 1702 TTTTGTAGACACAGGTTTTCGCATGTTGCCAGGCTGGTCTTGAATCCTGAGCTCAA 1761  
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||  
Db 22375 TTTTGTAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAG 22316  
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||  
QY 1762 GCAACCTCGCGGCTCGGCTCCCAAAGTACTGGGATTACACGAGAAGGCACCATGCC 1821  
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||  
Db 22315 GCAATCTGCCATCTCGGCTCCCAAAGTCTGGGATTACAAGCATGAGCCACCATGCC 22256  
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||  
QY 1822 AGCTAGATCT 1832  
|| || || || || || || || || || || || || || || || || || || || ||  
Db 22255 AGACTATTTAT 22245

Search completed: November 17, 2002, 16:01:28  
Job time : 4970 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 13:08:43 ; Search time 368 Seconds  
(without alignments)  
16363.705 Million cell updates/sec

Title: US-09-817-199A-1

Perfect score: 2674

Sequence: 1 ttgcctcgccggccgacct.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2581.6	96.5	2623	22	AAS27036
2	2240	83.8	2323	22	AAI60852
3	1286	48.1	1316	22	AAS27458
4	1099.6	41.1	1106	22	AAI59066
5	1079	40.4	1109	22	AAH75184
6	873	32.6	875	22	ABN83696
7	795.4	29.7	797	21	AAA52610
8	568	21.2	576	24	ABA02774
9	550.8	20.6	843	20	AAH86720

10	480	18.0	576	24	ABA02773	Mouse degranulation
11	468.6	17.5	475	23	AAS81431	DNA encoding novel
12	292.6	10.9	1340	19	AAV65197	Human RAB protein,
13	274	10.2	443	21	AAH97117	Mouse Exol04 nucle
14	237	8.9	2001	23	ABL29523	Drosophila melanog
15	209.4	7.8	964	22	AAS27040	cDNA encoding nove
16	207.4	7.8	24000	21	AAA88551	Human dual-specific
17	206.6	7.7	1786	22	AAH08354	Human secreted pro
18	204.8	7.7	10859	22	AAH04720	Human reproductive
19	204.8	7.7	10859	22	ABL97627	Human testicular a
20	203.4	7.6	138169	21	AAA34791	Human adenosine re
21	203.4	7.6	141589	21	AAF20913	Human ELAM-1 polyn
22	203.4	7.6	141589	21	AAF21127	Human low adenosin
23	203.4	7.6	141589	21	AAF21152	Human low adenosin
24	203.4	7.6	141589	21	AAA35005	Human adenosine re
25	203.4	7.6	141589	21	AAA35030	Human adenosine re
26	203.4	7.6	146981	21	AAF21442	Human ELAM-1 polyn
27	203.4	7.6	209273	21	AAF21437	Human factor-relat
28	203	7.6	409	22	AAK70203	Human immune/haema
29	202.8	7.6	27579	22	AAK66484	Human immune/haema
30	202.6	7.6	29222	24	ABL39412	Human electron-tra
31	202.4	7.6	10468	22	AAK77961	Human immune/haema
32	201.8	7.5	103747	24	ABQ88139	Human osteoblast d
33	201.6	7.5	52845	22	AAK71437	Human immune/haema
34	201.4	7.5	4969	22	ABA16169	Human nervous syst
35	201.4	7.5	6720	24	ABO88138	Human osteoblast d
36	201.2	7.5	33353	22	AAK70003	Human immune/haema
37	201	7.5	7428	22	ABA07164	Human pancreatic c
38	200.8	7.5	21835	22	AAK66483	Human immune/haema
39	200.6	7.5	21833	22	AAL35910	Human musculoskele
40	200.4	7.5	10663	22	ABA18961	Human nervous syst
41	200.4	7.5	31865	22	AAK85330	Human immune/haema
42	200.4	7.5	31865	22	AAK85331	Human immune/haema
43	200.2	7.5	21777	22	AAK74096	Human immune/haema
44	199.8	7.5	1130	22	ABA20629	Human nervous syst
45	199.8	7.5	1130	22	AAL37008	Human musculoskele

## ALIGNMENTS

RESULT 1

AAS27036

ID AAS27036 standard; cDNA; 2623 BP.

XX AAS27036;

XX AAS27036;

XX 07-NOV-2001 (first entry)

XX cDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;

XX Antinflammatory; anti-HIV; antibacterial; antinflammatory; cancer;

XX Immune system disorder; rheumatoid arthritis; inflammatory condition;

XX Organ transplant rejection; infection; hepatitis C; blood disorder;

XX Sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;

XX Neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

XX Chromosomal abnormality; Down syndrome; ischaemia; renal disorder;

XX Cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

XX Reproductive system; gastrointestinal; liver disorder; AIDS; ss;

XX Acquired immune deficiency syndrome.

XX Homo sapiens.

XX WO200154733-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-0501312.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

PR	02-MAR-2000;	2000US-0186350;
PR	16-MAR-2000;	2000US-0189076;
PR	17-MAR-2000;	2000US-0199074;
PR	18-APR-2000;	2000US-0198123;
PR	19-MAY-2000;	2000US-0205515;
PR	07-JUN-2000;	2000US-0209467;
PR	28-JUN-2000;	2000US-0214886;
PR	30-JUN-2000;	2000US-0215135;
PR	07-JUL-2000;	2000US-0216647;
PR	07-JUL-2000;	2000US-0216680;
PR	11-JUL-2000;	2000US-0217487;
PR	11-JUL-2000;	2000US-0217496;
PR	14-JUL-2000;	2000US-0218290;
PR	26-JUL-2000;	2000US-0220963;
PR	26-JUL-2000;	2000US-0220964;
PR	14-AUG-2000;	2000US-0224518;
PR	14-AUG-2000;	2000US-0224519;
PR	14-AUG-2000;	2000US-0225213;
PR	14-AUG-2000;	2000US-0225216;
PR	14-AUG-2000;	2000US-0225266;
PR	14-AUG-2000;	2000US-0225267;
PR	14-AUG-2000;	2000US-0225268;
PR	14-AUG-2000;	2000US-0225270;
PR	14-AUG-2000;	2000US-0225477;
PR	14-AUG-2000;	2000US-0225757;
PR	14-AUG-2000;	2000US-0225758;
PR	14-AUG-2000;	2000US-0225759;
PR	18-AUG-2000;	2000US-0226279;
PR	22-AUG-2000;	2000US-0226681;
PR	22-AUG-2000;	2000US-0226688;
PR	22-AUG-2000;	2000US-0229344;
PR	22-AUG-2000;	2000US-0229345;
PR	05-SEP-2000;	2000US-0229509;
PR	05-SEP-2000;	2000US-0229513;
PR	06-SEP-2000;	2000US-0230437;
PR	06-SEP-2000;	2000US-0230438;
PR	08-SEP-2000;	2000US-0231242;
PR	08-SEP-2000;	2000US-0231243;
PR	08-SEP-2000;	2000US-0231244;
PR	08-SEP-2000;	2000US-0231413;
PR	08-SEP-2000;	2000US-0231414;
PR	08-SEP-2000;	2000US-0232080;
PR	08-SEP-2000;	2000US-0232081;
PR	12-SEP-2000;	2000US-0231968;
PR	14-SEP-2000;	2000US-0232397;
PR	14-SEP-2000;	2000US-0233063;
PR	14-SEP-2000;	2000US-0233064;
PR	14-SEP-2000;	2000US-0233065;
PR	21-SEP-2000;	2000US-0234223;
PR	21-SEP-2000;	2000US-0234274;
PR	25-SEP-2000;	2000US-0234997;
PR	25-SEP-2000;	2000US-0234998;
PR	26-SEP-2000;	2000US-0235484;
PR	27-SEP-2000;	2000US-0235834;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0236327;
PR	29-SEP-2000;	2000US-0236367;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236369;
PR	29-SEP-2000;	2000US-0236370;
PR	02-OCT-2000;	2000US-0236602;
PR	02-OCT-2000;	2000US-0237037;
PR	02-OCT-2000;	2000US-0237038;
PR	02-OCT-2000;	2000US-0237039;

PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246533.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250161.
PR	01-DEC-2000;	2000US-0250390.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-)	HUMAN GENOME SCI INC
XX		
XX	Rosen CA,	Barash SC, Ruben
PI		
XX	WPI; 2001-465460/50.	
DR	P-PSDB; RAU17119.	
DR		
XX		
XX	Novel polypeptides useful for	
PT	prognosing disorders related	
PT	to disorders and neuronal disor	
XX		
XX	Claim 1; SEQ ID No 71.	880pp



XX The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative  
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
CC Addison's disease), reproductive system disorders, gastrointestinal  
CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
CC as stimulators of B-cell responsiveness to pathogens, activators of  
CC T-cells, to induce higher affinity antibodies, and as a means to induce  
CC tumour proliferation in pathologies e.g. acquired immune deficiency  
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
CC pathway protein coding sequences and PCR primers of the invention.  
XX

Query Match 96.5%; Score 2581.6; DB 22; Length 2623;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 42 ATGACGGGCACGCCAGGCGCGTTGCCACCGGATGGGAGGCCGCCAGCGCTCCCGG 101  
DB 12 ATGACGGGCACGCCAGGCGCGTTGCCACCGGATGGGAGGCCGCCAGCGCTCCCGG 71  
QY 102 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 161  
DB 72 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 131  
QY 162 GGCACCAATGTTCTGTATCAATTCACAAAGACGGGCGCTTCTGTCGGAACCTTCATCA 221  
DB 132 GGCACCAATGTTCTGTATCAATTCACAAAGACGGGCGCTTCTGTCGGAACCTTCATCA 191  
QY 222 GCCACCGTCGGCATAGACTTCAGGACAGGTGGTGACTGTGGATGGCGTGAGAGTGAG 281  
DB 192 GCCACCGTCGGCATAGACTTCAGGACAGGTGGTGACTGTGGATGGCGTGAGAGTGAG 251  
QY 282 CTGCAGATCTGGACACCGCTGGGAGGAACGGTTCGGAAGCGTCAACCATGCTATTAC 341  
DB 252 CTGCAGATCTGGACACCGCTGGGAGGAACGGTTCGGAAGCGTCAACCATGCTATTAC 311  
QY 342 AGAGATGCTCAGCGCTTGCTTCTGTGTATGACATCACCAACAAATCTTCTTCGACAC 401  
DB 312 AGAGATGCTCAGCGCTTGCTTCTGTGTATGACATCACCAACAAATCTTCTTCGACAC 371  
QY 402 ATCAGGCGCTGCTCACTAGATTATGATGATGCCAGAGGACGTGTGATCATGCTG 461  
DB 372 ATCAGGCGCTGCTCACTAGATTATGATGATGCCAGAGGACGTGTGATCATGCTG 431  
QY 462 CTAGGCAACAAGCGGATATGACAGCGAAAGATGATCCGTTCCGAAGCAGGAGACC 521  
DB 432 CTAGGCAACAAGCGGATATGACAGCGAAAGATGATCCGTTCCGAAGCAGGAGACC 491  
QY 522 TTGGCCAGGAGTACGGTGTTCCTTCTTGGAGACAGCGCCAAAGACTGGCATGAATGTG 581  
DB 492 TTGGCCAGGAGTACGGTGTTCCTTCTTGGAGACAGCGCCAAAGACTGGCATGAATGTG 551  
QY 582 GAGTTAGCCTTTCTGGCCATCCCAAGGAACCTGAATATCCGGCGGGCATCAGGCGGAT 641  
DB 552 GAGTTAGCCTTTCTGGCCATCCCAAGGAACCTGAATATCCGGCGGGCATCAGGCGGAT 611  
QY 642 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGGAAGCGCTCCAGCTGCTGC 701  
DB 612 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGGAAGCGCTCCAGCTGCTGC 671

QY 702 TCCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTTGAGGACACACAGGATGACGCCT 761  
DB 672 TCCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTTGAGGACACACAGGATGACGCCT 731  
QY 762 TCCCGCTCCAGCGCTTGCTTATCCAGAGGCTGAGCCAAATGGGAGAAAGATGAGGA 821  
DB 732 TCCCGCTCCAGCGCTTGCTTATCCAGAGGCTGAGCCAAATGGGAGAAAGATGAGGA 791  
QY 822 CTCACCTGACAGCGCTTCTTACGAGGAGCTATACCTCAACTCCTTACTTGTAGTTCCTGC 881  
DB 792 CTCACCTGACAGCGCTTCTTACGAGGAGCTATACCTCAACTCCTTACTTGTAGTTCCTGC 851  
QY 882 GGTCTCCCGCATPCCACAGAGGGTAAACACTTTAGCTTTTATTTAATAGTACATAAT 941  
DB 852 GGTCTCCCGCATPCCACAGAGGGTAAACACTTTAGCTTTTATTTAATAGTACATAAT 911  
QY 942 TTAATACCAAAAAGGCGCTGATCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCT 1001  
DB 912 TTAATACCAAAAAGGCGCTGATCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCT 971  
QY 1002 TTGCTTTTGTAGGACTTTGGGGCGCGCTTCTTCTTAAAGCATTAACAAAGGTGTTGC 1061  
DB 972 TTGCTTTTGTAGGACTTTGGGGCGCGCTTCTTCTTAAAGCATTAACAAAGGTGTTGC 1031  
QY 1062 TCAGCTCAGCCCGCAGGGGACAGATGACATTTGGGGGTGAGGGGAGGTAAATGACTCA 1121  
DB 1032 TCAGCTCAGCCCGCAGGGGACAGATGACATTTGGGGGTGAGGGGAGGTAAATGACTCA 1091  
QY 1122 TGCACCCCTCAGTTTACGCTGGAGAGGCTCAGGTACCCAGCCCTTACTGTCTCCCGC 1181  
DB 1092 TGCACCCCTCAGTTTACGCTGGAGAGGCTCAGGTACCCAGCCCTTACTGTCTCCCGC 1151  
QY 1182 TCTCCAGTGTATCTTCCGCCCATCTCCCAATTAAGTGGGCGCTTGTGCTGTAGGAA 1241  
DB 1152 TCTCCAGTGTATCTTCCGCCCATCTCCCAATTAAGTGGGCGCTTGTGCTGTAGGAA 1211  
QY 1242 GACCAAAAGCCTCAGGGAAGATAGAGATGAGAGTGGAGGGGAGGACAAAGGGGAGA 1301  
DB 1212 GACCAAAAGCCTCAGGGAAGATAGAGATGAGAGTGGAGGGGAGGACAAAGGGGAGA 1271  
QY 1302 GAGTAGGCTGTAGCTGGCTATCTTGGCGCTTACTTACA-CCCCCCTGGAGGATGCCCT 1360  
DB 1272 GAGTAGGCTGTAGCTGGCTATCTTGGCGCTTACTTACAACCCCCCTGGAGGATGCCCT 1331  
QY 1361 TTTCTCCAGCACACACAGCATTTGGGCGACCTGGAATATTGGTTCAGGCTCCTGTCT 1420  
DB 1332 TTTCTCCAGCACACACAGCATTTGGGCGACCTGGAATATTGGTTCAGGCTCCTGTCT 1391  
QY 1421 CTGGAGTTCAGATCTCTGGGGAGCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTTTC 1480  
DB 1392 CTGGAGTTCAGATCTCTGGGGAGCCCTCCCCCCCCCTGAATCCCTGGCTTAGCTTTC 1451  
QY 1481 CTGGCTGTGCACCTTAAACACCTCAGGTGAGACTAGGAAAGAGTTTGTATTTATTTT 1540  
DB 1452 CTGGCTGTGCACCTTAAACACCTCAGGTGAGACTAGGAAAGAGTTTGTATTTATTTT 1511  
QY 1541 TTGAATGTAGTCTCTTCTGTCGCCAGGCTGAGGTGAGTGAATTCCTCCGCTCAC 1600  
DB 1512 TTGAATGTAGTCTCTTCTGTCGCCAGGCTGAGGTGAGTGAATTCCTCCGCTCAC 1571  
QY 1601 TCAACCTTCACTCTCTGGGGTCAAGGATCTCCACCTCAGCGCGGAGTAGCTGG 1660  
DB 1572 TCAACCTTCACTCTCTGGGGTCAAGGATCTCCACCTCAGCGCGGAGTAGCTGG 1631  
QY 1661 GACTATAGTGTGTACCATCACACCTGGCTGAATTTTGTATTTTGTAGACACAGGTT 1720  
DB 1632 GACTATAGTGTGTACCATCACACCTGGCTGAATTTTGTATTTTGTAGACACAGGTT 1691  
QY 1721 TCGCCATGTTCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACTGCCGGCTCGGC 1780  
DB 1692 TCGCCATGTTCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACTGCCGGCTCGGC 1751  
QY 1781 CTCCCAAGTACTGGGATTACAGCGAAGGACCATGCCAGGCTAGATGTGTCTTATC 1840

|||||  
Db 1752 CTCCAAAGTACTGGATTACACGAGAGGACCAATGCCAGGCTAGATGCTCTATC 1811  
QY 1841 CCAATCTTTGGCAGGCTCAGCTCCACAGGCGATTCTTCAAGCAGCTGAAGTGTTTA 1900  
Db 1812 CCAATCTTTGGCAGGCTCAGCTCCACAGGCGATTCTTCAAGCAGCTGAAGTGTTTA 1871  
QY 1901 GCCTCTCTGGTTAAGAGCCAGATAGAGAAATCCCTTTCTTAGTGTGGAAATGTGTG 1960  
Db 1872 GCCTCTCTGGTTAAGAGCCAGATAGAGAAATCCCTTTCTTAGTGTGGAAATGTGTG 1931  
QY 1961 TG-AAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAGCAAAACC 2019  
Db 1932 TGAATAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAGCAAAACC 1991  
QY 2020 TCCCTCTGACATGATCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCAACCATGCCCT 2079  
Db 1992 TCCCTCTGACATGATCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCAACCATGCCCT 2051  
QY 2080 TTCTTTAACTTCTCAACAGATACCAGGCGCTAAACTGCTTTTACCTCCCTCTACTGA 2139  
Db 2052 TTCTTTAACTTCTCAACAGATACCAGGCGCTAAACTGCTTTTACCTCCCTCTACTGA 2111  
QY 2140 GTCAGGTAGTGGTGGAGGTCACCCATTTCCGAGTTTAAACCAATGCAATATGAGTAA 2199  
Db 2112 GTCAGGTAGTGGTGGAGGTCACCCATTTCCGAGTTTAAACCAATGCAATATGAGTAA 2171  
QY 2200 ACAAGTTCATGTGGTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCTCTCTCT 2259  
Db 2172 ACAAGTTCATGTGGTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCTCTCTCT 2231  
QY 2260 TGCTACATATCCCAAGCTCTGATTCCTGCTGCAATGGGAAGTGGAGAAACATGAGG 2319  
Db 2232 TGCTACATATCTCCCAAGCTCCGATCCCTGCTGCAATGGGAAGTGGAGAAACATGAGG 2291  
QY 2320 TCATGACCTGCAGGCTCTTACTGAGCTCTCCGCGCTGGAGGGGAGAGGGGAGGA 2379  
Db 2292 TCATGACCTGCAGGCTCTTACTGAGCTCTCCGCGCTGGAGGGGAGAGGGGAGGA 2351  
QY 2380 AGAAGTATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTC 2439  
Db 2352 AGAAGTATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTC 2411  
QY 2440 TCTGAGCTGACGCGCTCCAGTTGGGCGGATAGAGAGTTCTCCGTGGCCCTCCCTCA 2499  
Db 2412 TCTGAGCTGACGCGCTCCAGTTGGGCGGATAGAGAGTTCTCCGTGGCCCTCCCTCA 2471  
QY 2500 GGCAGAGCGAGGAGGCTGACATTCGAGTCTCTCTGGGCGCCCAAGCGAGTTGCGAG 2559  
Db 2472 GGCAGAGCGAGGAGGCTGACATTCGAGTCTCTCTGGGCGCCCAAGCGAGTTGCGAG 2531  
QY 2560 GAGATCCCAATCCCATAGACAGCTCTGGGCTCTTTGATTTTTCAGAAATTAACAT 2619  
Db 2532 GAGATCCCAATCCCATAGACAGCTCTGGGCTCTTTGATTTTTCAGAAATTAACAT 2591  
QY 2620 GCAGTATTTGGAAACCAAAAAA 2651  
Db 2592 GCAGTATTTGGAAACCAAAAAA 2623

RESULT 2  
AAI60852  
ID AAI60852 standard; cDNA; 2323 BP.  
XX  
AC AAI60852;  
DT  
XX 22-OCT-2001 (first entry)  
DE Human polynucleotide SEQ ID NO 4841.  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR P-PSDB; AAM41696.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Claim 1; SEQ ID NO 4841; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Query Match 83.8%; Score 2240; DB 22; Length 2323;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 2298; Conservative 0; Mismatches 16; Indels 9; Gaps 4;  
QY 15 GGCACCTGCTCAGCTCTCGTCCAGGACATGAGGGGACGCGCGCTGGCCACCCCG 74  
Db 1 GGAATTCCTCAGCTCTCGTCCAGGACATGAGGGGACGCGCGCTGGCCACCCCG 60  
QY 75 GATGGGAGGCGCGCGCGCTCCCGCGCTCCCGCGCTCCCGCGCTCCCGCGCTCCCGCG 134  
Db 61 GATGGGAGGCGCGCGCGCTCCCGCGCTCCCGCGCTCCCGCGCTCCCGCGCTCCCGCG 120  
QY 135 GTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCAAATCAAGAC 194  
Db 121 GTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCAAATCAAGAC 180  
QY 195 GGGGCTTCTGTCGGAACCTTCATAGCCACGCTCGGATAGACTTCAGGAACAGGTTG 254  
Db 181 GGGGCTTCTGTCGGAACCTTCATAGCCACGCTCGGATAGACTTCAGGAACAGGTTG 240

QY 255 GTGACTGTGGATGGCGTGAAGCTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGCGAGGAACGG 314  
DB 241 GTGACTGTGGATGGCGTGAAGCTGAAGCTGAGATCTGGGACACCGCTGGCGAGGAACGG 300  
QY 315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTGCTTCTGCTGTATGAC 374  
DB 301 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTGCTTCTGCTGTATGAC 360  
QY 375 ATCACCAACAAATCTCTTTTCACAAACATCAGGGCGCTGGCTCAGTCAGATTCATCAGTAT 434  
DB 361 ATCACCAACAAATCTCTTTTCACAAACATCAGGGCGCTGGCTCAGTCAGATTCATCAGTAT 420  
QY 435 GCCCAGAGGCGTGGTGATCATGCTGTAGCAACAAGGCGGATPATGAGCAGCGAAAGA 494  
DB 421 GCCCAGAGGCGTGGTGATCATGCTGTAGCAACAAGGCGGATPATGAGCAGCGAAAGA 480  
QY 495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTAGGGTTCCTTCCTGGAG 554  
DB 481 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTAGGGTTCCTTCCTGGAG 540  
QY 555 ACCAGCGCAAGACTGGCATGAATGTGGAGTTAGGCTTTCTGGCCATCGCCAAAGAACTG 614  
DB 541 ACCAGCGCAAGACTGGCATGAATGTGGAGTTAGGCTTTCTGGCCATCGCCAAAGAACTG 600  
QY 615 AAATACCGGGCGGCATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 674  
DB 601 AAATACCGGGCGGCATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 660  
QY 675 TCCCAAGAGAGCGTCCAGCTGCTGCTTCAATGTAATCCAGGGGCGAGAGAGAG 734  
DB 661 TCCCAAGAGAGCGTCCAGCTGCTGCTTCAATGTAATCCAGGGGCGAGAGAGAG 720  
QY 735 GCTCTGGAGGCACAGAGTGCAGCTTCCGCTCCAGGCTGGCTTATCCAGAGGC 794  
DB 721 GCTCTGGAGGCACAGAGTGCAGCTTCCGCTCCAGGCTGGCTTATCCAGAGGC 780  
QY 795 TGAGCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCTAGCAGGGAGCTA 854  
DB 781 TGAGCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCTAGCAGGGAGCTA 840  
QY 855 TACTCCAACCTCTACTGATTCCTGCGGTCTCCCGCATCCACAGGGAGGTAACAC 914  
DB 841 TACTCCAACCTCTACTGATTCCTGCGGTCTCCCGCATCCACAGGGAGGTAACAC 900  
QY 915 TTAGCTTTTATTTAATAGTACATAATTTAATACCAAAAAGGCGCTGGATCCCAAAA 974  
DB 901 TTAGCTTTTATTTAATAGTACATAATTTAATACCAAAAAGGCGCTGGATCCCAAAA 960  
QY 975 AACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTTTCTAGGACTTGGGGGCGCGCCCTCC 1034  
DB 961 AACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTTTCTAGGACTTGGGGGCGCGCCCTCC 1020  
QY 1035 TCCTAAGCATACAAGGTGGTGTGCTCCAGCTCAGCCCGAGGGGACACAGATGCAT 1094  
DB 1021 TCCTAAGCATACAAGGTGGTGTGCTCCAGCTCAGCCCGAGGGGACACAGATGCAT 1080  
QY 1095 TGGGGGTGAGGCGAGTATGACTCCATCGCACCTCAGTTTCACTGGACAGAGCTCAG 1154  
DB 1081 TGGGGGTGAGGCGAGTATGACTCCATCGCACCTCAGTTTCACTGGACAGAGCTCAG 1140  
QY 1155 GTGACCCCGAGCCTTCACTGTCTCCGCTCTCCAGGAGCTTATCTTCCGCCCATCTCCCAA 1214  
DB 1141 GTGACCCCGAGCCTTCACTGTCTCCGCTCTCCAGGAGCTTATCTTCCGCCCATCTCCCAA 1200  
QY 1215 ATAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGAGATATGA 1274  
DB 1201 ATAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGAGATATGA 1260  
QY 1275 GATGGAGGGGAGCACAGGGGACAGAGTAGGCTCTAGCTGGCTATCTCTGGCCCTTAC 1334  
DB 1261 GATGGAGGGGAGCACAGGGGACAGAGTAGGCTCTAGCTGGCTATCTCTGGCCCTTAC 1320

QY 1335 TAACACCCCGCTGGAGGATGCCCTTTTCTCCAGCACACAAGACATATGGGGCACTTGG 1394  
DB 1321 TAACACCCCGCTGGAGGATGCCCTTTTCTCCAGCACACAAGACATATGGGGCACTTGG 1380  
QY 1395 AAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGAGGCCCTCCCCC 1454  
DB 1381 AAATATTGGTTCCAGGCTCCTGTTCTCTGGACTTCAGATCCTGGGGAGGCCCTCCCCC 1440  
QY 1455 CCTGAATCCCTGGCTTAGCTACCTTCCCTGTGCACCTAAAAACCTCAGTCTCAGAAT 1514  
DB 1441 CCTGAATCCCTGGCTTAGCTACCTTCCCTGTGCACCTAAAAACCTCAGTCTCAGAAT 1500  
QY 1515 AGAAAAGAGTTTCTTTTATTTTAAATAGAGTCTCGTTCTGTGCGCCAGGCTGA 1574  
DB 1501 AGAAAAGAGTTTCTTTTATTTTAAATAGAGTCTCGTTCTGTGCGCCAGGCTGA 1560  
QY 1575 GGTGAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCT 1634  
DB 1561 GGTGAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCT 1620  
QY 1635 CCACCTCAGCGCGCGAAGTAGTGGGACTATAGTGTGTACCATCACACCTGGGCTAAAT 1694  
DB 1621 CCACCTCAGCGCGCGAAGTAGTGGGACTATAGTGTGTACCATCACACCTGGGCTAAAT 1680  
QY 1695 TTTGTATTTTGTAGACACAGGGTTTCGCCATGTTCGCCAGGCTGGTCTTGAATTCCTG 1754  
DB 1681 TTTGTATTTTGTAGACACAGGGTTTCGCCATGTTCGCCAGGCTGGTCTTGAATTCCTG 1740  
QY 1755 AGCTCAACCAACCTCGCGGCTCGGCTCCCAAGTAGTGGATTTACAGCGAAGGCAAC 1814  
DB 1741 AGCTCAACCAACCTCGCGGCTCGGCTCCCAAGTAGTGGATTTACAGCGAAGGCAAC 1800  
QY 1815 CATGCCAGGCTAGATGTCTTATCCCAATCTTTGGCAGGCATGCAGCTCCACAGGG 1874  
DB 1801 CATGCCAGGCTAGATGTCTTATCCCAATCTTTGGCAGGCATGCAGCTCCACAGGG 1860  
QY 1875 ATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCTCTGGGTTAAGAGCCAGATAAGGAGAA 1934  
DB 1861 ATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCTCTGGGTTAAGAGCCAGATAAGGAGAA 1920  
QY 1935 CCCTTCTCTAGTGTGGATGTGTGTAAGAAAAGAGAAATCCCTGGCTCCTGGAGCT 1994  
DB 1921 CCCTTCTCTAGTGTGGATGTGTGTAAGAAAAGAGAAATCCCTGGCTCCTGGAGCT 1980  
QY 1995 GGTGGAGACAAGATTAAGCAAAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTG 2054  
DB 1981 GGTGGAGACAAGATTAAGCAAAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTG 2040  
QY 2055 CCTCTCCTGTGACCACCCATGCCCTTCTTTAACTTCTCAAAACAGATACAGGGGCTAA 2114  
DB 2041 CCTCTCCTGTGACCACCCATGCCCTTCTTTAACTTCTCAAAACAGATACAGGGGCTAA 2100  
QY 2115 ACTGCTTTTACCTCCCTCTACTGAGTCAAGTTAGTGTGGGAGG---TCACCCATTTTC 2171  
DB 2101 ACTGCTTTTACCTCCCTCTACTGAGTCAAGTTAGTGTGGGAGGTTCAACCCATTTTC 2160  
QY 2172 CGAGTT---AAACCAATGCAATA-TGAGTAAAAAAGT---CATGTGGGTATGTCTGGGG 2225  
DB 2161 CGAGTTTAAACCAATGCNATATTGAGTAAACAAGATTCAATGTGGGGTATGTCTGGGG 2220  
QY 2226 TAGAGAGGGGTAGCAAGTTTCATGTCTCTCTTTGGTGCACATATCTCCCAAGCTCTGA 2285  
DB 2221 TAGAGAGGGGTAGCAAGTTTCATGTCTCTCTTTGGTGCACATATCTCCCAAGCTCTGA 2280  
QY 2286 TCCCTGCCATGGGAAGTGGACAGGAACATGAGGTCATGACCT 2328  
DB 2281 TCCCTNCCATGGGAAGTGGACAGGAACATGAGGTCATGACCT 2323

RESULT 3

AAS27458

ID AAS27458 standard; cDNA; 1316 BP.

XX

AC AAS27458;  
XX  
XX DT 07-NOV-2001 (first entry)  
XX DE cDNA encoding novel signal transduction pathway protein, Seq ID 493.  
XX  
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antineoplastic; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
PN  
PN WO200154733-A1.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01312.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 03-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 14-SEP-2000; 2000US-0233066.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX P-PSDB; RAU17541.  
XX WPI: 2001-465460/50.  
XX P-PSDB; RAU17541.  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
XX prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders -  
XX  
XX Claim 1; SEQ ID No 493; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
XX diagnosing, preventing and treating diseases including immune system  
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
XX transplant rejections and graft versus host disease, infectious diseases  
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
XX other blood-related disorders (sickle cell anaemia), myeloproliferative  
XX disorders, primary haematopoietic disorders, hyperproliferative  
XX disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
XX abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
XX disorders (e.g. glomerulonephritis), cardiovascular disorders  
XX (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
XX wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
XX Addison's disease), reproductive system disorders, gastrointestinal  
XX disorder (inflammatory disorders), liver disorders (cirrhosis),  
XX as stimulators of B-cell responsiveness to pathogens, activators of  
XX T-cells, to induce higher affinity antibodies, and as a means to induce  
XX tumour proliferation in pathologies e.g. acquired immune deficiency  
XX syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
XX pathway protein coding sequences and PCR primers of the invention.  
XX  
Query Match 48.1%; Score 1286; DB 22; Length 1316;  
Best Local Similarity 98.5%; Pred. No. 5.5e-23;  
Matches 1286; Conservative 9; Mismatches 11; Indels 0; Gaps 0;  
QY 45 ACGGCGACGCCAGGCGCGTTCGCCACCGGGATGCGAGGCCCGCCGAGCGCTCCCGGCC 104  
DB 1 ACGGCGACGCCAGGCGCGCTTCGCCACCGGGATGCGAGGCCCGCCGAGCGCTCCCGGCC 60  
QY 105 TGCAGTCGCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGC 164  
DB 61 TGCAGTCGCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGC 120  
QY 165 AAAACATGTTCTGATCCAAATCAAGACGGGGGCTTCCTGCTCCGGAACTTCATAGCC 224  
DB 121 AAAACATGTTCTGATCCAAATCAAGACGGGGGCTTCCTGCTCCGGAACTTCATAGCC 180  
QY 225 ACCGTCGCGATAGACTTCAGGACAGGTGGTGACTGTGGATGGCGTGAGAGTGAGCTG 284  
DB 181 ACCGTCGCGATAGACTTCAGGACAGGTGGTGACTGTGGATGGCGTGAGAGTGAGCTG 240

QY 285 CAGATCTGGACACCGCTGGCGAGGACGGTTCGGAAGCGTCACCCATGCTTATTACAGA 344  
DB 241 CAGATCTGGACACCGCTGGCGAGGACGGTTCGGAAGCGTCACCCATGCTTATTACAGA 300  
QY 345 GATGCTCAGGCGCTTCTGCTGTATGACATCACCACAAATCTTCTTCCACACATC 404  
DB 301 GATGCTCAGGCGCTTCTGCTGTATGACATCACCACAAATCTTCTTCCACACATC 360  
QY 405 AGGGCTGGCTCAGTATGAGATTCATGATGATGATGATGATGATGATGATGATGATGAT 464  
DB 361 AGGGCTGGCTCAGTATGAGATTCATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 465 GGCACAAAGCGGATATGAGCAGCGAAGAGTATCGGTTCCGAAAGAGGAGACCTTGG 524  
DB 421 GGCACAAAGCGGATATGAGCAGCGAAGAGTATCGGTTCCGAAAGAGGAGACCTTGG 480  
QY 525 GCCAGGAGTAGGCTGTTCCCTTCTGGAGACCGGCCAAGACTGGCATGAATGTGGAG 584  
DB 481 GCCAGGAGTAGGCTGTTCCCTTCTGGAGACCGGCCAAGACTGGCATGAATGTGGAG 540  
QY 585 TTAGCTTTCTGGCCATCGCCAAAGAACTGAATACCGGGCGGGCATCAGGCGGATGAG 644  
DB 541 TTAGCTTTCTGGCCATCGCCAAAGAACTGAATACCGGGCGGGCATCAGGCGGATGAG 600  
QY 645 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCGAAGAGCGCTCCAGCTGCTGCTCC 704  
DB 601 CCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCGAAGAGCGCTCCAGCTGCTGCTCC 660  
QY 705 TTCATGTGAATCCAGGGGGCAGAGAGGCTCTGGAGGACACAGGATGAGCGCTTCC 764  
DB 661 TTCATGTGAATCCAGGGGGCAGAGAGGCTCTGGAGGACACAGGATGAGCGCTTCC 720  
QY 765 CCTCCCGAGGCTGCTTATTCACAGAGCTGAGGCAATGGGAGAAACATGGAGGACTC 824  
DB 721 CCTCCCGAGGCTGCTTATTCACAGAGCTGAGGCAATGGGAGAAACATGGAGGACTC 780  
QY 825 ACTGCACAGCCCTCTCTAGCAGGAGCTATCTCAACTCTCTACTTGTAGTTCCTGCGT 884  
DB 781 ACTGCACAGCCCTCTCTAGCAGGAGCTATCTCAACTCTCTACTTGTAGTTCCTGCGT 840  
QY 885 CTCGCCGATCCACAGGAGGCTAAACACTTAGCTTTTATTATAGTACATAATTTA 944  
DB 841 CTCGCCGATCCACAGGAGGCTAAACACTTAGCTTTTATTATAGTACATAATTTA 900  
QY 945 ATACCAAAAGAGCGCTGATCCCAAAAGAGGCTGGGAGCTAGTGGCCCTTTTG 1004  
DB 901 ATACCAAAAGAGCGCTGATCCCAAAAGAGGCTGGGAGCTAGTGGCCCTTTTG 960  
QY 1005 CTTTCTAGGACTTGGGGGCGCGCCCTCCCTCCTTAAGCATAAACAAAGGTGCTGCTCC 1064  
DB 961 CTTTCTAGGACTTGGGGGCGCGCCCTCCCTCCTTAAGCATAAACAAAGGTGCTGCTCC 1020  
QY 1065 AGCTCAGTCCCGAGGAGGACACAGATGCACATTTGGGGTGAGGCGAGTAATGACTCCATCG 1124  
DB 1021 AGCTCAGTCCCGAGGAGGACACAGATGCACATTTGGGGTGAGGCGAGTAATGACTCCATCG 1080  
QY 1125 CACCTCTCAGTTCAGTTCGAGGAGGCTCAGGTGACCCCGAGCTTCACTGCTCCCGCTCT 1184  
DB 1081 CACCTCTCAGTTCAGTTCGAGGAGGCTCAGGTGACCCCGAGCTTCACTGCTCCCGCTCT 1140  
QY 1185 CCAGGAGCTTATCTTGGCGCCCATCTCCCAAAATGAAGTGGGCGCTTGTGCTGTAGGAGAC 1244  
DB 1141 CCAGGAGCTTATCTTGGCGCCCATCTCCCAAAATGAAGTGGGCGCTTGTGCTGTAGGAGAC 1200  
QY 1245 CAAAGCCTCAGGAGAGATAAGAGATATGAGATGGAGGGGGAGGACAGGCGCAGAGAG 1304  
DB 1201 CAAAGCCTCAGGAGAGATAAGAGATATGAGATGGAGGGGGAGGACAGGCGCAGAGAG 1260  
QY 1305 TAGGGTCTAGTGGCTTATCTCTGGCCTTACTTAACACCCCTCCCTGGAG 1350  
DB 1261 TAGGGTCTAGTGGCTTATCTCTGGCCTTACTTAACACCCCTCCCTGGAG 1306



## RESULT 5

AAH75184  
ID AAH75184 standard; DNA; 1109 BP.  
XX  
AC AAH75184;  
XX  
DT 13-NOV-2001 (first entry)  
XX  
DE Nucleotide sequence of human 32712 G-protein.  
XX  
KW Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;  
KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;  
KW adult respiratory distress syndrome; Goodpasture's syndrome;  
KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;  
KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;  
KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;  
KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;  
KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;  
KW chronic bacterial meningoencephalitis; multiple sclerosis;  
KW amytropic lateral sclerosis; stroke; Huntington's disease; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 124..699  
FT /tag= a "G-protein"  
FT /product= "G-protein"  
XX  
PN WO200164887-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 27-FEB-2001; 2001WO-US06292.  
XX  
XX 29-FEB-2000; 2000US-0185606.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX PA  
XX PI Meyers RA;  
XX  
XX WPI; 2001-550182/61.  
DR P-PSDB; AAG67156.  
XX  
XX Novel human small G-protein polypeptides and polynucleotides for  
PT treating lung disorders, liver disorders and brain disorders -  
PT  
XX  
XX Claim 2; Fig 26; 151pp; English.  
XX  
XX The present sequence encodes a human G-protein. The specification  
CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The  
CC G-protein polypeptides and polynucleotides are useful as a target for  
CC diagnosis and treatment of G-protein mediated or related disorders,  
CC and for identifying agonists and antagonists for diagnosis and  
CC treatment. They are useful for treating disorders of lung (e.g.  
CC congenital anomalies, pulmonary congestion, oedema, adult respiratory  
CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,  
CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver  
CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,  
CC Wilson's disease, autoimmune hepatitis and hepatic failure), and  
CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute  
CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic  
CC bacterial meningoencephalitis, multiple sclerosis, amytropic lateral  
CC sclerosis, stroke and Huntington's disease).  
XX  
SQ Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;  
Query Match 40.4%; Score 1079; DB 22; Length 1109;  
Best Local Similarity 99.5%; Pred. No. 2.2e-197;  
Matches 1082; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 37 GGACATGACGGGCGCGCGTGGCCACCGGGATGGCGAGGCCCGCGCGCT 96

Db 23 GCGTCCGGACGGGCGACGCCAGCGCGTGGCCACCGGGATGGCGAGGCCCGCGCGCT 82  
QY 97 CCCCAGGCTGCGAGTCCGAGCTAGGACCTCAGGGGCAAGGTGATGCTTCTGGGAGACACAG 156  
Db 83 CCCCAGGCTGCGAGTCCGAGCTAGGACCTCAGGGGCAAGGTGATGCTTCTGGGAGACACAG 142  
QY 157 GCGTCCGGCAAAACATGTTTCTCTGATCCAAATCAAAAGACGGGGCTTCTCTGTCGGGAACCT 216  
Db 143 GCGTCCGGCAAAACATGTTTCTCTGATCCAAATCAAAAGACGGGGCTTCTCTGTCGGGAACCT 202  
QY 217 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTACTGTGGATGGCGTGAGAG 276  
Db 203 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTACTGTGGATGGCGTGAGAG 262  
QY 277 TGAAGCTGAGATCTGGGACACCGCTGGGCGAGGAACGGTTCGGAAGCGTCACCCATGCTT 336  
Db 263 TGAAGCTGAGATCTGGGACACCGCTGGGCGAGGAACGGTTCGGAAGCGTCACCCATGCTT 322  
QY 337 ATTACAGAGATGCTCAGGCGCTTCTCTGCTGATCAGATCACCACAATCTTCTTCG 396  
Db 323 ATTACAGAGATGCTCAGGCGCTTCTCTGCTGATCAGATCACCACAATCTTCTTCG 382  
QY 397 ACAACATCAGGCGCTGGCTCAGTATGATGATTCATGAGATTCATCCAGAGGGAGCGTGGTGATCA 456  
Db 383 ACAACATCAGGCGCTGGCTCAGTATGATGATTCATGAGTATGCCAGAGGGAGCGTGGTGATCA 442  
QY 457 TCGTCTAGGCAACAAGGGCGGATATGAGCAGCGGAAGAGATGATCCGTCGGAAGAGCGGAG 516  
Db 443 TCGTCTAGGCAACAAGGGCGGATATGAGCAGCGGAAGAGATGATCCGTCGGAAGAGCGGAG 502  
QY 517 AGACCTTGGCCAGGAGTACGGTGTTCCTCTGGAGACCGAGCGCCAGACAGCTGGCATGA 576  
Db 503 AGACCTTGGCCAGGAGTACGGTGTTCCTCTGGAGACCGAGCGCCAGACAGCTGGCATGA 562  
QY 577 ATGTGAGTATAGCTTCTGGCCATCGCAAGGAATGAAATACCGGGCGGCGCATCAGG 636  
Db 563 ATGTGAGTATAGCTTCTGGCCATCGCAAGGAATGAAATACCGGGCGGCGCATCAGG 622  
QY 637 CGGATCAGGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGAGCGCTCCAGCT 696  
Db 623 CGGATCAGGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGAGCGCTCCAGCT 682  
QY 697 GCTGCTCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATGC 756  
Db 683 GCTGCTCTTCATGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATGC 742  
QY 757 AGCCTTCCCTCCAGGCTGGCTTATTCGAAGAGGCTGAGCCAATGGGGAGAGAAAGATG 816  
Db 743 AGCCTTCCCTCCAGGCTGGCTTATTCGAAGAGGCTGAGCCAATGGGGAGAGAAAGATG 802  
QY 817 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATCTCAACTCCTACTTGAATG 876  
Db 803 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATCTCAACTCCTACTTGAATG 862  
QY 877 CTTGCGGCTCTCCCGCATCCACAGGAGGGTAAACACTTAGCTTTTATTAATAGTAC 936  
Db 863 CTTGCGGCTCTCCCGCATCCACAGGAGGGTAAACACTTAGCTTTTATTAATAGTAC 922  
QY 937 ATAATTTAATACAAAAAGGCGCTTGGATCCCAAAAAACCGAGGCTGGGAGCTAGTGG 996  
Db 923 ATAATTTAATACAAAAAGGCGCTTGGATCCCAAAAAACCGAGGCTGGGAGCTAGTGG 982  
QY 997 CCGTTTGTCTTCTAGACTTGGGGGGCGGCCCTCCCTCTAAGCATATCAAGGTGGT 1056  
Db 983 CCGTTTGTCTTCTAGACTTGGGGGGCGGCCCTCCCTCTAAGCATATCAAGGTGGT 1042  
QY 1057 GTTGCTCCAGCTCAGCCCGAGGGGACACAGATGCACTTTGGGGGTGAGGCGAGTATGA 1116  
Db 1043 GTTGCTCCAGCTCAGCCCGAGGGGACACAGATGCACTTTGGGGGTGAGGCGAGTATGA 1102  
QY 1117 CTCATC 1123  
|||||







OS Homo sapiens.  
XX WO200032630-A2.  
XX 08-JUN-2000.  
XX 06-DEC-1999; 99WO-US28773.  
XX 04-DEC-1998; 98US-01111006.  
XX (SEAR) SEARLE & CO G D.  
XX Dotson SB, Ma X;  
XX WPI; 2000-465041/40.  
XX Novel nucleic acids derived from activated eosinophil cells useful for  
XX treating allergic diseases such as asthma comprises a specific  
XX nucleotide sequence -  
XX  
XX Claim 1; Page 95; 125pp; English.  
XX The present invention relates to a number of nucleotide sequences  
XX which encode proteins involved in the activation of eosinophils  
XX (AA52601-A52679). Eosinophils are involved in immune reactions, and  
XX these genes and their proteins provide possible targets for new drugs to  
XX combat diseases such as asthma, allergic rhinitis, atopic dermatitis,  
XX anaphylaxis, allergic bronchopulmonary aspergillosis, eczema, psoriasis,  
XX emphysema, leukaemia, lymphomas, ovarian cancer, pneumonia, immune  
XX disorders, some connective tissue disorders, and inflammatory conditions  
XX including septic shock, arthritis, nephritis, inflammatory bowel disease  
XX and Crohn's disease.  
XX Sequence 797 BP; 192 A; 201 C; 209 G; 195 T; 0 other;  
SQ  
Query Match 29.7%; Score 795.4; DB 21; Length 797;  
Best Local Similarity 99.9%; Pred. No. 3.9e-143;  
Matches 796; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1840 CCCAATCCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGT 1899  
DB 1 CCCAATCCCTTTGGCAGGCATGCAGCTCCACAGGCGATTTCTTCAAGCAGCTGCAGTGT 60  
QY 1900 AGCCCTCTCTGGTGTAAAGCCAGATGAAGGAGAAATCCCTTCTAGTGTGGAATGTGT 1959  
DB 61 AGCCCTCTCTGGTGTAAAGCCAGATGAAGGAGAAATCCCTTCTAGTGTGGAATGTGT 120  
QY 1960 GTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAAGCAAAACC 2019  
DB 121 GTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAAGCAAAACC 180  
QY 2020 TCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCTCCCTGACCACCATGCCCT 2079  
DB 181 TCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCCTCTCCCTGACCACCATGCCCT 240  
QY 2080 TTCTTTAACTTCTCAACAGATACACAGGCGCTTAACTGCTTTACCTCCCTCTACTCA 2139  
DB 241 TTCTTTAACTTCTCAACAGATACACAGGCGCTTAACTGCTTTACCTCCCTCTACTCA 300  
QY 2140 GTCAGTTAGTGGTGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGAGTAA 2199  
DB 301 GTCAGTTAGTGGTGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGAGTAA 360  
QY 2200 ACAAGTCAATGGTATGCTCTGGGTAGAGAGGGGTAGCAAGTTATGTCCTCT 2259  
DB 361 ACAAGTCAATGGTATGCTCTGGGTAGAGAGGGGTAGCAAGTTATGTCCTCT 420  
QY 2260 TGCTCATATCTCCCAAGCTCTGATCCCTGCCATGGGAGTGGAGCAAGCAATGAGG 2319  
DB 421 TGCTCATATCTCCCAAGCTCTGATCCCTGCCATGGGAGTGGAGCAAGCAATGAGG 480  
QY 2320 TCATGACCTGCAGGCATCTTACTGTCAGCTCTGCCGCTGGAGGGGAGAGGGGAGGA 2379  
|||||

DB 481 TCATGACCTGCAGGCATCTTTACTGCAGCTCTGCCGCTGAGGGGAGAGGGGAGGA 540  
QY 2380 AGAAGTATCGCTGCACATTTCTGAGGCTACTGCAATTTCTTCAAGGAGCAAAATCTTGC 2439  
DB 541 AGAAGTATCGCTGCACATTTCTGAGGCTACTGCAATTTCTTCAAGGAGCAAAATCTTGC 600  
QY 2440 TCTGAGCAGTCAAGCGCTCCAGTTTGGGCCCGGATAGGAAGTTCTCCGTGGCCTCCCTCA 2499  
DB 601 TCTGAGCAGTCAAGCGCTCCAGTTTGGGCCCGGATAGGAAGTTCTCCGTGGCCTCCCTCA 660  
QY 2500 GGCAGAGCAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGCCCAAGCAGGTTGCAG 2559  
DB 661 GGCAGAGCAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGCCCAAGCAGGTTGCAG 720  
QY 2560 GAGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAAC 2619  
DB 721 GAGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAAC 780  
QY 2620 GCAGTATTTTGGAAAGC 2636  
DB 781 GCAGTATTTTGGAAAGC 797  
RESULT 8  
ABA02774  
ID ABA02774 standard; DNA; 576 BP.  
XX ABA02774;  
AC ABA02774;  
XX  
DT 07-FEB-2002 (first entry)  
XX  
DE Human degranulation regulator encoding DNA SEQ ID NO 3.  
XX  
KW Degranulation; mast cell; human; mouse; antiallergic; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..576  
FT */\*tag= a*  
FT */product= "degranulation regulator"*  
XX  
PN WO200179478-A1.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-JP03268.  
XX  
PR 19-APR-2000; 2000JP-0118408.  
XX  
PA (DAIN) DAINIPPON PHARM CO LTD.  
XX  
PI Yamada T, Ido M;  
XX  
DR WPI; 2002-041335/05.  
XX  
DR P-PSDB; AAM52190.  
XX  
PT Mast cell degranulation controller for treatment of allergies -  
XX  
PS Claim 11; Page 64-65; 85pp; Japanese.  
XX  
CC The invention relates to a protein for regulating degranulation of mast  
XX cells (degranulation regulators) and the encoding polynucleotides, with  
XX antiallergic activity, used in the treatment of allergies associated  
XX with degranulation of mast cells.  
XX  
SQ Sequence 576 BP; 138 A; 147 C; 168 G; 123 T; 0 other;  
Query Match 21.2%; Score 568; DB 24; Length 576;  
Best Local Similarity 99.1%; Pred. No. 1.2e-99;  
Matches 571; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 138 ATGCTTCTGGGAGACAGCGTCGCGCAAAACATGTTTCTGTATCAATTCAAAGACGG 197

Db 1 ATGCTTCTGGAGCTGGCGCTGGCAAAACATGTTTCTGTATCCAAATTCAAAGACGGG 60  
QY 198 GCCTTCCTGTCGGAACTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGTGGTG 257  
Db 61 GCCTTCCTGTCGGAACTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGTGGTG 120  
QY 258 ACTGTGATGGGTGAGAGTGAAGCTGAGATCTGGACACCGCTGGCAGGAACGGTTC 317  
Db 121 ACTGTGATGGGTGAGAGTGAAGCTGAGATCTGGACACCGCTGGCAGGAACGGTTC 180  
QY 318 CGAAGCTCACCATCTTATTACAGAGATGCTCAGGCTTGTCTGTGTATGACATC 377  
Db 181 CGAAGCTCACCATCTTATTACAGAGATGCTCAGGCTTGTCTGTGTATGACATC 240  
QY 378 ACCAACAATCTCTTCGCAACATCAGGCTTGGCTCACTGAGATTCATGATGATGCC 437  
Db 241 ACCAACAATCTCTTCGCAACATCAGGCTTGGCTCACTGAGATTCATGATGATGCC 300  
QY 438 CAGAGGACGTGTGATGATCTGCTAGGCAACAAGCGGATATGACGCGAAAGATG 497  
Db 301 CAGAGGACGTGTGATGATCTGCTAGGCAACAAGCGGATATGACGCGAAAGATG 360  
QY 498 ATCGTTCCCAAGCAGGAGACCTTGGCCAGGAGTACGGTGTTCCTTCCTGGAGACC 557  
Db 361 ATCGTTCCCAAGCAGGAGACCTTGGCCAGGAGTACGGTGTTCCTTCCTGGAGACC 420  
QY 558 AGCGCCAAAGACTGGCATGATGTGGAGTGTAGCCTTCTGGCCATCGCAAGGAACGAAA 617  
Db 421 AGCGCCAAAGACTGGCATGATGTGGAGTGTAGCCTTCTGGCCATCGCAAGGAACGAAA 480  
QY 618 TACCGGCGCGGCATCAGGCGGATGAGCCAGGTTCAGATCCAGACTATGATGAGTCC 677  
Db 481 TACCGGCGCGGCATCAGGCGGATGAGCCAGGTTCAGATCCAGACTATGATGAGTCC 540  
QY 678 CAGAAAGCGCTCCAGCTGCTCTCTTCATGTGA 713  
Db 541 CAGAAAGCGCTCCAGCTGCTCTCTTCATGTGA 576

RESULT 9  
AA86720  
ID AAX86720 standard; DNA; 843 BP.

XX AC AAX86720;

XX 27-OCT-1999 (first entry)

DE DNA sequence encoding an exocytotic protein designated Exo2.

XX Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;  
KW Inflammation; ss.

XX OS Mus sp.

XX FH Key Location/Qualifiers  
XX CDS 124..699

FT FT /\*tag= a  
FT FT /transl\_except= (pos: 136..138, aa: xaa)  
FT FT /transl\_except= (pos: 202..204, aa: xaa)  
FT FT /note= "Xaa is an unspecified amino acid"

XX PN WO9942586-A2.

XX PD 26-AUG-1999.

XX PF 23-FEB-1999; 99WO-US03909.

XX PR 26-MAY-1998; 98US-0086650.

XX PR 23-FEB-1998; 98US-0075534.

XX PA (RIGE-) RIGEL PHARM INC.

XX

PI Anderson D, Fisher J, Huang B, Lorens J, Luo Y;  
PI Shen M;  
XX  
DR WPI; 1999-518605/43.  
DR P-PSDB; AAY30132.  
XX  
PT New exocytotic proteins useful for diagnosis and treatment of  
PT exocytosis-mediated conditions and in drug screening  
XX  
PS Claim 6; Fig 6; 53pp; English.

XX The present sequence encodes an exocytotic protein designated Exo2.  
CC The protein has an inhibitory effect on exocytosis (i.e. the fusion of  
CC secretory vesicles with the cellular plasma membrane) and can be  
CC administered therapeutically to treat or prevent exocytosis-mediated  
CC disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving  
CC undesirable release of compounds via exocytosis e.g. inflammation  
CC mediated by the release compounds such as histamine. Exo2 is useful  
CC diagnostically and to produce antibodies useful to purify the proteins  
CC and therapeutically to reduce or eliminate the biological activity of  
CC the protein. Exo2 nucleic acids can be used therapeutically to increase  
CC Exo2 activity in cells by known gene therapy techniques. They can also  
CC be used to produce probes or primers to isolate Exo2 proteins from other  
CC organisms, especially humans. The nucleic acids, host cells and proteins  
CC are useful in screening assays to identify binding agents, especially  
CC drug screening assays to identify agonists and antagonists useful  
CC therapeutically to enhance or reduce Exo2 activity.

XX Sequence 843 BP; 195 A; 218 C; 239 G; 175 T; 16 other;

Query Match 20.6%; Score 550.8; DB 20; Length 843;  
Best Local Similarity 84.9%; Pred. No. 2.5e-96;  
Matches 635; Conservative 0; Mismatches 110; Indels 3; Gaps 2;

QY 32 GTCACAGGACATCAGCGGCACGCCAGGCGC--CGTTGCCACCCCGGATGCGAGGCCCC 89  
Db 16 GNNCGGNGANATGATGGCACACAGGAGCTTGTACCGCTTGGGATGCGAGGCCCT 75

QY 90 GAGCGCTCCCGCCCTCGAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCGGGA 149  
Db 76 GAGCGCTCCCGCCCTCAGCGGCAAGTACGATNTACCAGGAGTGTGCTTCCTTGA 135

QY 150 GACACAGGCGTCGGCAAAACATGTTCCCTGATCCATTCAGGACGCGGCTTCCTGTCC 209  
Db 136 GANTCGGCGCTCGGCAAAACCTGTTCCCTGATCCATTCAGGACGCGGCTTCCTGTCC 195

QY 210 GGAACCTTCATAGCCACCGTCGCGCATAGACTTCAGGAACAAGGTGTGATGTGGATGC 269  
Db 196 GGAACCTTCATAGCCACCGTCGCGCATAGACTTCAGGAACAAGGTGTGATGTGGATGC 255

QY 270 GTGAGAGTGAAGTGCAGATCTGGGACACCGCTGGGAGGAGGAGTTCGGAAGCGTACC 329  
Db 256 TCCAGGGGTGAAGCTTCAGATCTGGGACATGCGAGGACGAGGCGCTTCGCGAGTGAAC 315

QY 330 CATGCTTATTACAGAGATGCTCAGGCTTCTGCTGTGTATGATCATCACAACAATCT 389  
Db 316 CATGCTTATTACAGAGATGCTCAGGCTTCTGCTGTGTATGATCATCACAACAAGTTC 375

QY 390 TCTTTTCACAACATCAGGCGCTGCTCAGTACGATTCATGAGTATGCCAGAGGAGCGTG 449  
Db 376 TCTTTTCACAACATCAGGCGCTGCTCAGTACGATTCATGAGTATGCCAGAGGAGCGTG 435

QY 450 GTGATCATGCTGTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGTCCGTCGGAA 509  
Db 436 GTGATCATGCTGTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTGTCCGTCGGAA 495

QY 510 GACGAGAGACCTTGGCCAGGAGTACGGTGTTCCTCTTCTGGAGACCGGCAAGACT 569  
Db 496 GATGGAGAGACACTGGCCAGGGAATATGTTCTTCTTCATGGAGACCACTGCGCAAGACT 555

QY 570 GGCATGAATGTGGAGTGTAGCTTCTGCGCATCGCCAAAGAACTGAAATACCGGCGGG 629  
Db 556 GGCATGAATGTGGAGTGTAGCTTCTGCGCAATGCGCAAGAACTGAAATACCGGCGGG 615

QY	630	CATPCAGCGCGATGAGCCCAAGCTTCCAGATCCGAGACTATGTAGTCCAGAGAAGCGC	689			
Db	616	AGGCAGCCTGATGAGCCCAAGCTTCCAGATCCGAGACTATGTGGAGTCCCGAAGAAGCGC	675			
QY	690	TCCAGCTGCTGCTCCCTTCATGTGTAATCCACAGGGGCGAGAGAGGCTCTGCGAGGCAC-A	748			
Db	676	TCCAGCTGCTGCTCCCTTCGTGTGACCCCTAGGGGCTAAGAGAGGGCCCAANAGACCCTT	735			
QY	749	CAGGATCCAGCCCTTCCCTCCAGGCC	776			
Db	736	GGGGATCGAGTACTCCCACTGCCACACC	763			
RESULT 10						
ABAO2773						
ID	ABAO2773 standard; DNA; 576 BP.					
AC	ABAO2773;					
XX						
DT	07-FEB-2002 (first entry)					
DE	Mouse degranulation regulator encoding DNA SEQ ID NO 2.					
XX	Degranulation; mast cell; human; mouse; antiallergic; ds.					
KW						
XX	Mus musculus.					
OS						
Key	Location/Qualifiers					
FT	1..576					
CDS	/*tag= a					
FT	/transl_except= (pos:247..249,aa:Lys)					
FT	/product= "degranulation regulator"					
XX						
PN	W0200179478-A1.					
XX						
PD	25-OCT-2001.					
XX						
PF	16-APR-2001; 2001WO-JP03268.					
XX						
PR	19-APR-2000; 2000JP-0118408.					
XX						
PA	(DAIN ) DAINIPPON PHARM CO LTD.					
XX						
PI	Yamada T, Ido M;					
XX						
DR	WPI; 2002-041335/05.					
DR	P-PSDB; AAM52189.					
XX						
PT	Mast cell degranulation controller for treatment of allergies -					
XX						
PS	Claim 10; Page 63-64; 85pp; Japanese.					
XX						
CC	The invention relates to a protein for regulating degranulation of mast cells (degranulation regulators) and the encoding polynucleotides, with antiallergic activity, used in the treatment of allergies associated with degranulation of mast cells.					
XX						
SQ	Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;					
Query Match 18.0%; Score 480; DB 24; Length 576;						
Best Local Similarity 89.6%; Pred. No. 8.5e-83;						
Matches 516; Conservative 0; Mismatches 60; Indels 0; Gaps						
QY	138	ATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCAATTCAAAGACGGG	197			
Db	1	ATGCTTCTGGGAGACTCGGCGCTCGGCAAAACCTGTTTCTGATCAATTCAAAGACGGG	60			
QY	198	GCCTTCTCTGTCGCGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAGGTGGTG	257			
Db	61	GCCTTCTCTGTCGCGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTG	120			
QY	258	ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCCGCTGGGACGAGTTC	317			

CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;

Query Match 17.5%; Score 468.6; DB 23; Length 475;

Best Local Similarity 99.2%; Pred. No. 1.3e-80;

Matches 471; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GGCACGTCTCACTCTCGTCCAGGACATGACGGGACGCGCGGTGGCCACCCCGG 74  
DB 1 GGAATTCTCTCACTCTCGTCCAGGACATGACGGGACGCGCGGTGGCCACCCCGG 60  
QY 75 GATGGGAGGCCCCCGAGCGCTCCCGCGCTCGAGTCCGAGCTACGACCTCAGGGCGAAG 134  
DB 61 GATGGGAGGCCCCCGAGCGCTCCCGCGCTCGAGTCCGAGCTACGACCTCAGGGCGAAG 120  
QY 135 GTGATGCTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATTCAAAGAC 194  
DB 121 GTGATGCTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATTCAAAGAC 180  
QY 195 GGGGCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254  
DB 181 GGGGCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240  
QY 255 GTCACTGTGATGGCTGAGAGTGAAGTGCAGATCTGGGACACCGCTGGCAGGACCG 314  
DB 241 GTCACTGTGATGGCTGAGAGTGAAGTGCAGATCTGGGACACCGCTGGCAGGACCG 300  
QY 315 TTCGGAAGCGTCACTGCTTTATACAGAGATGCTCAGGCCCTGCTTCTGCTGTATGAC 374  
DB 301 TTCGGAAGCGTCACTGCTTTATACAGAGATGCTCAGGCCCTGCTTCTGCTGTATGAC 360  
QY 375 ATCACCACAAATCTCTTTTCGACACATCAGGGCGCTGGCTCACTGAGATTCATGAGTAT 434  
DB 361 ATCACCACAAATCTCTTTTCGACACATCAGGGCGCTGGCTCACTGAGATTCATGAGTAT 420  
QY 435 GCCCAGAGGACGCTGTGATCATGCTGCTAGGCAACAAGCCGATATGACGACCG 489  
DB 421 GCCCAGAGGACGCTGTGATCATGCTGCTAGGCAACAAGCCGATATGACGACCG 475

RESULT 12

AAV65197

ID AAV65197 standard; DNA; 1340 BP.

XX AC AAV65197;

XX 17-DEC-1998 (first entry)

XX Human RAB protein, SRAB, coding sequence.

XX Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;

XX Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;

XX viral infection; therapy; autoimmune sialosis; cystic fibrosis;

XX diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;

XX hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;

polycystic renal disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 43..615

FT /tag= a

FT /product= "SRAB"

XX W09842839-A1.

XX 01-OCT-1998.

XX 25-MAR-1998; 98WO-US03871.

XX 26-MAR-1997; 97US-0824873.

XX (INCY-) INCYTE PHARM INC.

XX Guegler KJ, Hillman JL;

XX WPI; 1998-542282/46.

XX P-PSDB; AAW80747.

XX New human Rab protein, SRAB - useful for treating disorders

XX associated with SRAB expression, including vesicle trafficking,

XX viral infection, and cancer

XX Claim 5; Fig 1; 31pp; English.

XX This sequence encodes the human RAB protein, SRAB, of the invention. The

XX SRAB protein can be used in compositions useful in the diagnosis,

XX prevention, or treatment of disorders associated with vesicle trafficking

XX (including epilepsy, Huntington's disease, Parkinson's disease and

XX schizophrenia), cancer, or viral infection. Host cells containing the DNA

XX sequence can be used to produce SRAB recombinantly. The viral infections

XX that can be treated using SRAB include those caused by retroviruses

XX particularly HIV and HTLV, hepatitisviruses, particularly hepatitis C,

XX hantaviruses, herpesviruses, and arboviruses. The cancers that can be

XX treated include adenocarcinoma, leukaemia, lymphoma, melanoma and

XX sarcoma, particularly cancers of the endocrine, gastrointestinal and

XX nervous systems and cancers of the adrenal gland, brain, breast, colon,

XX oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,

XX prostate, salivary gland, stomach, thyroid, and uterus. The protein can

XX also be used to treat conditions such as autoimmune sialosis,

XX choroideaemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,

XX hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme

XX deficiency, pancreatitis, and polycystic renal disease.

XX Query Match 10.9%; Score 292.6; DB 19; Length 1340;

XX Best Local Similarity 70.9%; Pred. No. 6.7e-47;

XX Matches 416; Conservative 0; Mismatches 169; Indels 2; Gaps 2;

QY 116 CTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175

DB 21 CTACGACGTCGCTTCAAGGTGATGCTTGTGGGGGACGCGGTGGGGAAGACCTGTCT 80

QY 176 CTTGA-TCCAAATCAAGACGGGCGCTTCTGTCCGGACCTTCATAGCCACGTCGGCA 234

DB 81 GCTGGGTGCGATTCAAGGATGGTGTCTTCTGGCGGGGACCTTCATCTCCACCGT-AGCA 139

QY 235 TAGACTTCAGGAACAAGGTGGTGACTGTGTGAGTGGAGTGAAGCTGCAGATCTGGG 294

DB 140 TTGACTTCGCGAACAAGTTCCTGGACGTGGATGGTGTGAAGTGAAGCTGCAGATCTGGG 199

QY 295 ACACCCCTGGGAGGACCGTTCCGAAGCTCACCCATGCTTATTACAGAGATGCTCAGG 354

DB 200 ACACAGCTGGTCAGGAGCGGTTCGCGAGTGTACCCATGCTACTACCGGGATGCTCATG 259

QY 355 CTTGCTTCTGCTGTATGACATCACCAAAATCTTCTTCGACACATCAGGGCGCTGGC 414

Sequence 443 BP: 109 A: 111 C: 131 G: 90 T: 2 other: 22

Best Local Similarity 84.00; PECO: NO: 2.1E-45;  
 Matches 307; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

RESULT 14  
ABL29523  
ID ABL29523 standard: DNA: 2001 BP

XX DE CC-MA 9-0004 Cont. p. 1

XX

OS *Drosophila melanogaster*.

XX  
37-SEP-2001

XX









Result No.	Score	Query Match	Length	DB	ID	Description
1	2674	100.0	2674	10	US-09-817-199A-1	Sequence 1, Appl
2	2031.8	76.0	13182	10	US-09-817-199A-3	Sequence 3, Appl
3	1086	40.6	1116	10	US-09-794-257-13	Sequence 13, Appl
4	873	32.6	875	12	US-10-051-986-10	Sequence 10, Appl
5	683.4	25.6	24707	10	US-09-740-027-3	Sequence 3, Appl
6	600.6	22.5	601	10	US-09-817-199A-24	Sequence 24, Appl
7	576	21.5	576	601	US-09-794-257-15	Sequence 15, Appl
8	573	21.4	601	10	US-09-817-199A-25	Sequence 25, Appl
9	443.4	16.6	447	10	US-09-867-550-1811	Sequence 1811, Ap
c 10	200.6	7.5	21833	10	US-09-764-877-2275	Sequence 2275, Ap
c 11	199.8	7.5	1130	10	US-09-764-877-3373	Sequence 3373, Ap
c 12	198.6	7.4	15745	10	US-09-764-877-3236	Sequence 3236, Ap
c 13	198.6	7.4	17216	10	US-09-764-877-3565	Sequence 3565, Ap
c 14	198.6	7.4	17217	10	US-09-764-877-3566	Sequence 3566, Ap
c 15	198.4	7.4	4962	10	US-09-764-868-1302	Sequence 1302, Ap
c 16	198.4	7.4	4962	10	US-09-764-869-1903	Sequence 1903, Ap
c 17	197.6	7.4	7791	10	US-09-764-877-3700	Sequence 3700, Ap
c 18	197.2	7.4	2754	10	US-09-745-763-133	Sequence 133, Appl
c 19	197.2	7.4	15515	10	US-09-822-860-3	Sequence 3, Appl

Db	241	TCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCG	300
Qy	301	CTGGCAGGAACCGTTCCGAAGCGTCACCCATGCTTTATTACAGAGATGCTCAGGCGTTGC	360
Db	301	CTGGCAGGAACCGTTCCGAAGCGTCACCCATGCTTTATTACAGAGATGCTCAGGCGTTGC	360
Qy	361	TTCTGTGTATGACATACACAACAAATCTTTTCGACAACATCAGGCGCTGGTCACTG	420
Db	361	TTCTGTGTATGACATACACAACAAATCTTTTCGACAACATCAGGCGCTGGTCACTG	420
Qy	421	AGATTCATGAGTATGCCAGAGGGAGCTGGTGATCATGCTGCTAGGCAACAGGCGGATA	480
Db	421	AGATTCATGAGTATGCCAGAGGGAGCTGGTGATCATGCTGCTAGGCAACAGGCGGATA	480
Qy	481	TGACACGGAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGAGTAGCGTG	540
Db	481	TGACACGGAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGAGTAGCGTG	540
Qy	541	TTCCCTTCTGGAGACAGCGCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCA	600
Db	541	TTCCCTTCTGGAGACAGCGCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCA	600
Qy	601	TCGCCAAGGAACCTGAATACCGGGCGGGCATCAGCGCGGATGACCCAGCTTCCAGATCC	660
Db	601	TCGCCAAGGAACCTGAATACCGGGCGGGCATCAGCGCGGATGACCCAGCTTCCAGATCC	660
Qy	661	GAGACTATGTAGAGTCCCAGAAGACGCTCCAGCTGCTGCTCTTCATGTGAATCCCAAG	720
Db	661	GAGACTATGTAGAGTCCCAGAAGACGCTCCAGCTGCTGCTCTTCATGTGAATCCCAAG	720
Qy	721	GGGCAGAGAGAGGCTCTGGAGGCACACAGGATGCAGCTTCCCGCTCCACAGGCTTGC	780
Db	721	GGGCAGAGAGAGGCTCTGGAGGCACACAGGATGCAGCTTCCCGCTCCACAGGCTTGC	780
Qy	781	TTATTCACAGAGCTCAGCCCAATCGGGAGAAAGATGGAGGACTCAGTCACAGCGCGTTC	840
Db	781	TTATTCACAGAGCTCAGCCCAATCGGGAGAAAGATGGAGGACTCAGTCACAGCGCGTTC	840
Qy	841	CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCGCATCCACAG	900
Db	841	CTAGCAGGGAGCTATACTCCAACTCCTACTTGAGTTCCTGCGGTCTCCCGCATCCACAG	900
Qy	901	GGAGGGTAAACACTTAGCTTTTATTAAATAGTACATAATTAATACAAAAAGGCGC	960
Db	901	GGAGGGTAAACACTTAGCTTTTATTAAATAGTACATAATTAATACAAAAAGGCGC	960
Qy	961	CTGATCCCCAAAAACCGAGCTGGAGCTAGTGGCGCTTTTGTCTTCAGGACTTTGGG	1020
Db	961	CTGATCCCCAAAAACCGAGCTGGAGCTAGTGGCGCTTTTGTCTTCAGGACTTTGGG	1020
Qy	1021	GGGCGGCGCTCCCTCCTAAGCATACAAGGTGGTGTGCTCCAGCTCAGCGCCACAGGG	1080
Db	1021	GGGCGGCGCTCCCTCCTAAGCATACAAGGTGGTGTGCTCCAGCTCAGCGCCACAGGG	1080
Qy	1081	ACACAGATGACATTTGGGGTGAGGCGAGGTAATGACTCCATTCGACCGCTCAGTTCAGCT	1140
Db	1081	ACACAGATGACATTTGGGGTGAGGCGAGGTAATGACTCCATTCGACCGCTCAGTTCAGCT	1140
Qy	1141	GGACAGGCTCAGGTGACCCAGGCTTCATGCTCTCCGCTCTCCAGGAGCTTATCTTC	1200
Db	1141	GGACAGGCTCAGGTGACCCAGGCTTCATGCTCTCCGCTCTCCAGGAGCTTATCTTC	1200
Qy	1201	GCOCCTATCCCCAAATAAGTGGGCGCTTGTGCTGTGAGGAAGACCAAGGCTCAGGGAAG	1260
Db	1201	GCOCCTATCCCCAAATAAGTGGGCGCTTGTGCTGTGAGGAAGACCAAGGCTCAGGGAAG	1260
Qy	1261	ATAAGAGATATGGAGATGGAGGGGGAGACAAGGGGCAGAGTAGGGTCTAGCTGGCT	1320
Db	1261	ATAAGAGATATGGAGATGGAGGGGGAGACAAGGGGCAGAGTAGGGTCTAGCTGGCT	1320
Qy	1321	ATCTCTGGGCTTACTAACACCCCTTGAGGAGTGCCTTTTCTCCAGCACAAGCAC	1380
Db	1321	ATCTCTGGGCTTACTAACACCCCTTGAGGAGTGCCTTTTCTCCAGCACAAGCAC	1380

Qy	1381	ATTGGGACCTGGAATAATGTGTTCCAGGCTCCTGTTCTCTGGAACTTTCAGATCCTCGGG	1441
Db	1381	ATTGGGACCTGGAATAATGTGTTCCAGGCTCCTGTTCTCTGGAACTTTCAGATCCTCGGG	1440
Qy	1441	GAGCCCTCCCCCTGAACTCCTGGCTTAGCTACCTTCTGCTGCTGTGCACCTTAAAC	1500
Db	1441	GAGCCCTCCCCCTGAACTCCTGGCTTAGCTACCTTCTGCTGCTGTGCACCTTAAAC	1500
Qy	1501	CTCAGGTCAGAACTAGGAAAGAGTTTGTGTTTTATTTTTTGAATGGAGTCTCGTTCT	1560
Db	1501	CTCAGGTCAGAACTAGGAAAGAGTTTGTGTTTTATTTTTTGAATGGAGTCTCGTTCT	1560
Qy	1561	GTGCCCAGGCTGAGTGCAGTAGTGCATCTCCGTCACATACCACTTCCACTCCCTGGG	1620
Db	1561	GTGCCCAGGCTGAGTGCAGTAGTGCATCTCCGTCACATACCACTTCCACTCCCTGGG	1620
Qy	1621	GCTCAAGCGATCTCCACCTCAGCGCCGCAAGTAGCTAGCTGGGACTATAGTGTGTACCATC	1680
Db	1621	GCTCAAGCGATCTCCACCTCAGCGCCGCAAGTAGCTAGCTGGGACTATAGTGTGTACCATC	1680
Qy	1681	ACACTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTGCCCAGGCTG	1740
Db	1681	ACACTGGCTAATTTTTGTATTTTTTGTAGACACAGGGTTTCGCCATGTGCCCAGGCTG	1740
Qy	1741	GTCTTGAATTCCTGAGCTCAAGCAACTGCCGGCTCGGCTCCCAAAGTAGTCTGGGATTA	1800
Db	1741	GTCTTGAATTCCTGAGCTCAAGCAACTGCCGGCTCGGCTCCCAAAGTAGTCTGGGATTA	1800
Qy	1801	CACGAGAAGCCACCATGCCAGGCTAGATGTGTTATCCCAATCTTTGGCAGGCATG	1860
Db	1801	CACGAGAAGCCACCATGCCAGGCTAGATGTGTTATCCCAATCTTTGGCAGGCATG	1860
Qy	1861	CAGCTCCACAGGGGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGACC	1920
Db	1861	CAGCTCCACAGGGGATTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCCTGGGTTAAGACC	1920
Qy	1921	AGATAAGGAGAAATCCCTTCTTAGTGTTTGGAATGTGTTGAAAAAAGAGAAATCCC	1980
Db	1921	AGATAAGGAGAAATCCCTTCTTAGTGTTTGGAATGTGTTGAAAAAAGAGAAATCCC	1980
Qy	1981	TGGCTCTGAGCTGTGGGAGACAAAGATTAGCAAAACCTCCCTGCATGTATCCCTTT	2040
Db	1981	TGGCTCTGAGCTGTGGGAGACAAAGATTAGCAAAACCTCCCTGCATGTATCCCTTT	2040
Qy	2041	GACCCCAAGCTCTGCCTCCTCCCTGACCAACCATGCCCTTCCCTTAACTTCTCAAAACAG	2100
Db	2041	GACCCCAAGCTCTGCCTCCTCCCTGACCAACCATGCCCTTCCCTTAACTTCTCAAAACAG	2100
Qy	2101	ATACCGAGGCTAACTGCTTTACCTCCCTCCTACTGAGTCAGGTTAGTGTGGGAGG	2160
Db	2101	ATACCGAGGCTAACTGCTTTACCTCCCTCCTACTGAGTCAGGTTAGTGTGGGAGG	2160
Qy	2161	TCACCCATTTCCGAGTTAAACCAATAGTATAGTAAAAAAGTCATGTGGGTATGTC	2220
Db	2161	TCACCCATTTCCGAGTTAAACCAATAGTATAGTAAAAAAGTCATGTGGGTATGTC	2220
Qy	2221	TGGGTTAGAGAGAGGGGTAGCAAGTTTCATGTCTCCTCTGGTCACATATCTCCCAAAGC	2280
Db	2221	TGGGTTAGAGAGAGGGGTAGCAAGTTTCATGTCTCCTCTGGTCACATATCTCCCAAAGC	2280
Qy	2281	TCTGATCCCTGCCATGGGAAGTGGACAGAAACATGAGGTCATGACCTGCAGGCATCTTT	2340
Db	2281	TCTGATCCCTGCCATGGGAAGTGGACAGAAACATGAGGTCATGACCTGCAGGCATCTTT	2340
Qy	2341	ACTGCAGCTCTGCCGGCTCGGAGGGGAGAGGGGAGGAAGAAGTATGCGCTGCACATTT	2400
Db	2341	ACTGCAGCTCTGCCGGCTCGGAGGGGAGAGGGGAGGAAGAAGTATGCGCTGCACATTT	2400
Qy	2401	CTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTACAGGGCTCCA	2460
Db	2401	CTGAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTACAGGGCTCCA	2460

QY 2461 GTTTGGGCCCCAATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGACGAGGGAGGCTG 2520  
Db 2461 GTTTGGGCCCCAATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGACGAGGGAGGCTG 2520  
QY 2521 ACATTGCCAGTCTCTTCTGGGGCCCCAAGCAGGTTGCGAGGATCCAATCCATAGACAG 2580  
Db 2521 ACATTGCCAGTCTCTTCTGGGGCCCCAAGCAGGTTGCGAGGATCCAATCCATAGACAG 2580  
QY 2581 CTCCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAACCTGCAATGCAATTTTGGAAACAAAA 2640  
Db 2581 CTCCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAACCTGCAATGCAATTTTGGAAACAAAA 2640  
QY 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674  
Db 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

RESULT 2  
US-09-817-199A-3  
; Sequence 3, Application US/09817199A  
; Patent No. US20020142380A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001187  
; CURRENT APPLICATION NUMBER: US/09/817,199A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 13182  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(13182)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-817-199A-3

Query Match 76.0%; Score 2031.8; DB 10; Length 13182;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2039; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 589 CTTTCTGCCCATCCCAAGGAACCTGAATACCGGGCGGGCATCAGCGGATGAGCCCA 648  
Db 10060 CCATTGTCTCTTCTCAAGGGAACTGAATACCGGGCGGGCATCAGCGGATGAGCCCA 10119  
QY 649 GCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAACGCTCCAGCTGCTGCTCTTCA 708  
Db 10120 GCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAACGCTCCAGCTGCTGCTCTTCA 10179  
QY 709 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATGACGCTTCCCGCT 768  
Db 10180 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATGACGCTTCCCGCT 10239  
QY 769 CCCAGGCTGGCTTATTCACAGAGGCTGAGCCAAATGGGGAGAAAGATGGAGACTCACTG 828  
Db 10240 CCCAGGCTGGCTTATTCACAGAGGCTGAGCCAAATGGGGAGAAAGATGGAGACTCACTG 10299  
QY 829 CACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTCTACTTGAGTTCTCGGCTCTCC 888  
Db 10300 CACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTCTACTTGAGTTCTCGGCTCTCC 10359  
QY 889 CCGCATCCACAGGGAGGTTAAACACTTAGCTTTTATTTAATAGTACATAATTAATAC 948  
Db 10360 CCGCATCCACAGGGAGGTTAAACACTTAGCTTTTATTTAATAGTACATAATTAATAC 10419  
QY 949 CAAAAAGCGCTGGATCCCAAAAAACCGAGGCTGGAGGCTAGTGGCCCTTTTGGCTTT 1008  
Db 10420 CAAAAAGCGCTGGATCCCAAAAAACCGAGGCTGGAGGCTAGTGGCCCTTTTGGCTTT 10479

QY 1009 CTAGGACTTGGGGGCGGSCCTCCCTCCTAAGCATAAACAAAGTGGTGTGCTCCAGCT 1068  
Db 10480 CTAGGACTTGGGGGCGGSCCTCCCTCCTAAGCATAAACAAAGTGGTGTGCTCCAGCT 10539  
QY 1069 CAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAAATGACTCCTCCGACC 1128  
Db 10540 CAGCCCCAGGGGACACAGATGCACTTTGGGGGTGAGGGCAGGTAAATGACTCCTCCGACC 10599  
QY 1129 CTCAGTTGAGTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGTCCTCCGCTCTCCAG 1188  
Db 10600 CTCAGTTGAGTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGTCCTCCGCTCTCCAG 10659  
QY 1189 GAGCTTATCTTCGCCCCATCTCCCAATAAGTGGGCCCTTGTCTGTGAGGAGACCAAA 1248  
Db 10660 GAGCTTATCTTCGCCCCATCTCCCAATAAGTGGGCCCTTGTCTGTGAGGAGACCAAA 10719  
QY 1249 GCCTCAGGGAAGATAGAGATATGGAGATGGAGGGGAGGACAGGGGCGAGAGTAGG 1308  
Db 10720 GCCTCAGGGAAGATAGAGATATGGAGATGGAGGGGAGGACAGGGGCGAGAGTAGG 10779  
QY 1309 GTCTAGCTGGCTATCTCTGGCCCTTACTAACACCCCTCGAGGSCATGCCCTTTCTCCA 1368  
Db 10780 GTCTAGCTGGCTATCTCTGGCCCTTACTAACACCCCTCGAGGSCATGCCCTTTCTCCA 10839  
QY 1369 GCACACAGCACAATTGGGGCACCTGGAATAATTGTTCCAGGCTCCTGTTCTCTGGACTT 1428  
Db 10840 GCACACAGCACAATTGGGGCACCTGGAATAATTGTTCCAGGCTCCTGTTCTCTGGACTT 10899  
QY 1429 CAGATCTGGGGAGGCCCTCCCCCCCCTGAATCCCTGGCTTACCTTCTGCTGT 1488  
Db 10900 CAGATCTGGGGAGGCCCTCCCCCCCCTGAATCCCTGGCTTACCTTCTGCTGT 10959  
QY 1489 GCACCTAAAAAAGCTCAGGTCAGAACTAGAGAAAGAGTTTGTATTTTGTGAAATG 1548  
Db 10960 GCACCTAAAAAAGCTCAGGTCAGAACTAGAGAAAGAGTTTGTATTTTGTGAAATG 11019  
QY 1549 GAGTCTGCTTCTGTGCGCCAGGCTCAGGTGCAAGTGTGCAATCTCCGCTCACTACAACT 1608  
Db 11020 GAGTCTGCTTCTGTGCGCCAGGCTCAGGTGCAAGTGTGCAATCTCCGCTCACTACAACT 11079  
QY 1609 CCACCTCTGGGGCTCAAGCGATCTCCACCTCAGCCCGCCGAGTAGCTGGGACTATAG 1668  
Db 11080 CCACCTCTGGGGCTCAAGCGATCTCCACCTCAGCCCGCCGAGTAGCTGGGACTATAG 11139  
QY 1669 GTGTGTACCATCACACTCGCTAATTTTGTATTTTGTAGACACAGGTTTTCGCCATG 1728  
Db 11140 GTGTGTACCATCACACTCGCTAATTTTGTATTTTGTAGACACAGGTTTTCGCCATG 11199  
QY 1729 TTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCTCGGCTCCCAAA 1788  
Db 11200 TTGCCCAGGCTGGTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCTCGGCTCCCAAA 11259  
QY 1789 GTACTGGGATACAGCAGAAAGCCATGCCCAGGCTAGATGTGCTTATGCCAATCT 1848  
Db 11260 GTACTGGGATACAGCAGAAAGCCATGCCCAGGCTAGATGTGCTTATGCCAATCT 11319  
QY 1849 TTGGCAGCATCGCTCCACAGGCTTCTTCAAGAGCTCAAGTGTGTTAGCCCTCT 1908  
Db 11320 TTGGCAGCATCGCTCCACAGGCTTCTTCAAGAGCTCAAGTGTGTTAGCCCTCT 11379  
QY 1909 GGGTTAAAGCCAGATAGGAGAAATCCCTTCTAGGTTTGAATGTGTGTGAAAAAA 1968  
Db 11380 GGGTTAAAGCCAGATAGGAGAAATCCCTTCTAGGTTTGAATGTGTGTGAAAAAA 11439  
QY 1969 AAGAGAAATCTCTGGCTCTGAGTGTGGGAGACAGATTAAAGCAAACTCCCTCTGAC 2028  
Db 11440 AAGAGAAATCTCTGGCTCTGAGTGTGGGAGACAGATTAAAGCAAACTCCCTCTGAC 11499  
QY 2029 ATGTATCCCTTTGACCCCAAGCTCTGCCCTCTCCCTGACCACTGCTTCTTAA 2088  
Db 11500 ATGTATCCCTTTGACCCCAAGCTCTGCCCTCTCCCTGACCACTGCTTCTTAA 11559  
QY 2089 CTTCTCAACAGATACAGGGGCTTAACTGCTTTTACCTCCCTCTCTACTGAGTCAGTTA 2148

```
Db 11560 CTTCTCAACAGATACCGGGCCCTAAACCTGCTTTACCTCCCTCTCTACTGAGTCAGGTTA 11619
QY 2149 GGTGGTGGGAGGTCAACCCATTTCCGAGTTAAACCAATGCAATATGATAGTAAACAAAGTCA 2208
Db 11620 GGTGGTGGGAGGTCAACCCATTTCCGAGTTAAACCAATGCAATATGATAGTAAACAAAGTCA 11679
QY 2209 TGTGGGTATGCTGGGGTATAGAGAGGGGTAGCAAGTTCAATGTGCTCCTCCTTGGTGCACAT 2268
Db 11680 TGTGGGTATGCTGGGGTATAGAGAGGGGTAGCAAGTTCAATGTGCTCCTCCTTGGTGCACAT 11739
QY 2269 ATCTCCCAAGCTCTCATCCCTGCCATGGAAGTGGACAGGAACATGAGTGCATGACCT 2328
Db 11740 ATCTCCCAAGCTCTCATCCCTGCCATGGAAGTGGACAGGAACATGAGTGCATGACCT 11799
QY 2329 GCAGGCGATCTTTACTGCAGCTCTCGCGGCTCGAGGGGAGAGGGGAGGAAGTATG 2388
Db 11800 GCAGGCGATCTTTACTGCAGCTCTCGCGGCTCGAGGGGAGAGGGGAGGAAGTATG 11859
QY 2389 CGTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGCAGAAATCTTGCTCTGAGCAG 2448
Db 11860 CGTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGCAGAAATCTTGCTCTGAGCAG 11919
QY 2449 TCAGCGGCTCCAGTTTGGGCCGATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGCA 2508
Db 11920 TCAGCGGCTCCAGTTTGGGCCGATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGCA 11979
QY 2509 GGGAGGAGGCTGACATTTGCCAGTCTCTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAA 2568
Db 11980 GGGAGGAGGCTGACATTTGCCAGTCTCTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAA 12039
QY 2569 TCCCATAGACAGCTCGGGCTCTTGCATTTGAGTTTTCAGAACTGAGCATTT 2628
Db 12040 TCCCATAGACAGCTCGGGCTCTTGCATTTGAGTTTTCAGAACTGAGCATTT 12099
QY 2629 TGGAAAGCAA 2639
Db 12100 TGGAAAGCACA 12110

RESULT 3
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)...(699)
US-09-794-257-13

Query Match 40.6%; Score 1086; DB 10; Length 1116;
Best Local Similarity 99.5%; Pred. No. 1.6e-181;
Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 37 GGGACATGAGCGGCACCGCAGCGCGTGGCCACCGGGATGCGAGGCCCGCGAGCGGT 96
Db 23 GCGTCGGGACGGCGCAGCGCGCGTGGCCACCGGGATGCGAGGCCCGCGAGCGGT 82
QY 97 CCCCGCCCTGCATCCGAGCTACGACCTCACGGGCAAGGTGATGCTCTCTGGGAGACACAG 156
```

```
Db 83 CCCCGCCCTGCACTCCGAGCTACGACCTCACGGGCAAGGTGATGCTCTCTGGGAGACACAG 142
QY 157 GGTGGGCAAAAACATGTTTCTCTGATCCTCAATTTCAAAGACGGGCTTCTCTGTCCGGAACCT 216
Db 143 GGTGGGCAAAAACATGTTTCTCTGATCCTCAATTTCAAAGACGGGCTTCTCTGTCCGGAACCT 202
QY 217 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCGTGAGAG 276
Db 203 TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCGTGAGAG 262
QY 277 TGAAGCTGCAGATCTGGGACACCGCTGGGACGAGGTTCCGAAGCGTCAACCATGCTT 336
Db 263 TGAAGCTGCAGATCTGGGACACCGCTGGGACGAGGTTCCGAAGCGTCAACCATGCTT 322
QY 337 ATTACAGAGATGCTCAGGCTTGTCTCTGTATGACATCACCAACAAATCTCTTTTCG 396
Db 323 ATTACAGAGATGCTCAGGCTTGTCTCTGTATGACATCACCAACAAATCTCTTTTCG 382
QY 397 ACAACATCAGGCTGGCTGCTCAGTATGATGATTCATGATGATGCCAGAGGAGGTGGTATCA 456
Db 383 ACAACATCAGGCTGGCTGCTCAGTATGATGATTCATGATGATGCCAGAGGAGGTGGTATCA 442
QY 457 TGTGCTAGGCAACAGGCGGATATGAGCAGGAGAAAGTGTCCCTTCCGAAGACGGAG 516
Db 443 TGTGCTAGGCAACAGGCGGATATGAGCAGGAGAAAGTGTCCCTTCCGAAGACGGAG 502
QY 517 AGACCTTGGCCAGGGAGTACGGTGTTCCTTCTGTGAGACACAGCGCCAAAGACTGGCATGA 576
Db 503 AGACCTTGGCCAGGGAGTACGGTGTTCCTTCTGTGAGACACAGCGCCAAAGACTGGCATGA 562
QY 577 ATGTGAGTTAGCTTCTTGGCCATCGCAAGAACTGAAATACCGGGCGGGGCATCAGG 636
Db 563 ATGTGAGTTAGCTTCTTGGCCATCGCAAGAACTGAAATACCGGGCGGGGCATCAGG 622
QY 637 CGSATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGAGCGCTCCAGCT 696
Db 623 CGSATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGAGCGCTCCAGCT 682
QY 697 GCTGCTCTTCTATGTAATCCAGGGGGCAGAGAGGCTCTTGGAGGCACACAGATGC 756
Db 683 GCTGCTCTTCTATGTAATCCAGGGGGCAGAGAGGCTCTTGGAGGCACACAGATGC 742
QY 757 AGCTTCCCTCCCTCCAGGCTGGCTTATTCGAAGAGGCTGAGCCAATGGGAGAAAGATG 816
Db 743 AGCTTCCCTCCCTCCAGGCTGGCTTATTCGAAGAGGCTGAGCCAATGGGAGAAAGATG 802
QY 817 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGGCTATATCTCAACTCCTACTTCTGAT 876
Db 803 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGGCTATATCTCAACTCCTACTTCTGAT 862
QY 877 CCTGCGGTCTCCCGCATCCACAGGGAGGTAAACACTTAGCTTTTATTTAATAGTAC 936
Db 863 CCTGCGGTCTCCCGCATCCACAGGGAGGTAAACACTTAGCTTTTATTTAATAGTAC 922
QY 937 ATAAATTAATACAAAAAGGCGCTTGATCCCAAAAAACCGAGGCTGGGAGCTAGTG 996
Db 923 ATAAATTAATACAAAAAGGCGCTTGATCCCAAAAAACCGAGGCTGGGAGCTAGTG 982
QY 997 CCCTTTTGTCTTAGGACTTGGGGGGCGGCTTCCCTCTAAGCATACAAAGGTGGT 1056
Db 983 CCCTTTTGTCTTAGGACTTGGGGGGCGGCTTCCCTCTAAGCATACAAAGGTGGT 1042
QY 1057 GTTGCTCCAGCTCAGCCCGGAGGAGACACAGATGCACTTTGGGGGTGAGGCGAGGTAAATGA 1116
Db 1043 GTTGCTCCAGCTCAGCCCGGAGGAGACACAGATGCACTTTGGGGGTGAGGCGAGGTAAATGA 1102
QY 1117 CTCCTATCGACCT 1130
Db 1103 CTCCTATCGACCT 1116

RESULT 4
```

Query Match	32.6%	Score 873	DB 12	Length 875
Best Local Similarity	100.0%	Pred. No. 2.4e-144		
Matches 873	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	46	CGSGCAGCCAGCGCGCGTTGCGACCCGGGATGCGAGGCCCCCGAGCGCTCCCGCGCCT	105	
DB	1	CGSGCAGCCAGCGCGCGTTGCCACCCGGGATGCGAGGCCCCCGAGCGCTCCCGCGCCT	60	
QY	106	GCAGTCCGAGCTACGGACCTCACGGGCAAGTGATGCTTCTGGGAGACACAGGGCGTCGGCA	165	
DB	61	GCAGTCCGAGCTACGACCTCACGGGCAAGTGATGCTTCTGGGAGACACAGGGCGTCGGCA	120	
QY	166	AAACATGTTTCTGATCCAATTCAAAGCAGGGGCTTCCTGTCGGGAACCTTCATAGCCA	225	
DB	121	AAACATGTTTCTGATCCAATTCAAAGCAGGGGCTTCCTGTCGGGAACCTTCATAGCCA	180	
QY	226	CCGTGCGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC	285	
DB	181	CCGTGCGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGAGTGAAGCTGC	240	

Query Match	25.6%	Score 683.4	DB 10	Length 24707
Best Local Similarity	99.9%	Pred: No. 4.9e-11		
Matches 684	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1955	GTCTTGTGAAAAAAGAGAAATCCCTGCTCTCTGGAGCTGTGGGAGACAAGATTAAAG	2014	
Db	1	GTCTTGTGAAAAAAGAGAAATCCCTGCTCTCTGGAGCTGTGGGAGACAAGATTAAAG	60	

QY 2015 AAACCTCCCTGACATGATATCCCTTTGACCCCAAGCTCTGCTCCCTGACCAACCCAT 2074  
Db 61 AAACCTCCCTGACATGATATCCCTTTGACCCCAAGCTCTGCTCCCTGACCAACCCAT 120  
QY 2075 GCCCTTTCCCTTTAACTTCTCAACAGATACAGGGCTAAACTGCTTTTAACTCCCTCCCT 2134  
Db 121 GCCCTTTCCCTTTAACTTCTCAACAGATACAGGGCTAAACTGCTTTTAACTCCCTCCCT 180  
QY 2135 ACTGAGTCAGTTAGTGTGGAGGTCAACCATTTCCGAGTTAAACCAATGCAATATGA 2194  
Db 181 ACTGAGTCAGTTAGTGTGGAGGTCAACCATTTCCGAGTTAAACCAATGCAATATGA 240  
QY 2195 GTAACAAAGTCAATGTGGGTATGCTGGGTAGAGAGGGGTAGCAATGTCATGTGTC 2254  
Db 241 GTAACAAAGTCAATGTGGGTATGCTGGGTAGAGAGGGGTAGCAATGTCATGTGTC 300  
QY 2255 CTCCTTGGTCACATATCTCCCAAGCTCTGATCCCTGCAATGGAAGTGCACAGGAACA 2314  
Db 301 CTCCTTGGTCACATATCTCCCAAGCTCTGATCCCTGCAATGGAAGTGCACAGGAACA 360  
QY 2315 TGAGGTCAATGACCTGACGGCATCTTTACTGCAGCTCTGCCGGCTTGAGGGGGAGAGGGG 2374  
Db 361 TGAGGTCAATGACCTGACGGCATCTTTACTGCAGCTCTGCCGGCTTGAGGGGGAGAGGGG 420  
QY 2375 GAGGAAGAAGTATGCGGTGCACATTTCTGAGGTACTGCAATTTGCTTTCAAGGCAGAAAT 2434  
Db 421 GAGGAAGAAGTATGCGGTGCACATTTCTGAGGTACTGCAATTTGCTTTCAAGGCAGAAAT 480  
QY 2435 CTTGCTCTGAGCAGTCAGCGCTCCAGTTTGGCCGATAGGAAGTCTCCCTGGCCCTC 2494  
Db 481 CTTGCTCTGAGCAGTCAGCGCTCCAGTTTGGCCGATAGGAAGTCTCCCTGGCCCTC 540  
QY 2495 CCTCAGCAGCAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGGCCCAAGGAGGT 2554  
Db 541 CCTCAGCAGCAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGGCCCAAGGAGGT 600  
QY 2555 TGACAGGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAAT 2614  
Db 601 TGACAGGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAAT 660  
QY 2615 AAACCTCAGTATTTTGGAAAGCAA 2639  
Db 661 AAACCTCAGTATTTTGGAAAGCACA 685

RESULT 6  
US-09-817-199a-24  
; Sequence 24, Application US/09817199a  
; Patent No. US20020142380A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO01187  
; CURRENT APPLICATION NUMBER: US/09/817,199a  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-817-199a-24

Query Match 22.5%; Score 600.6; DB 10; Length 601;  
Best Local Similarity 99.8%; Pred. No. 8.6e-97;  
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1983 GCTCCTGGAGCTGGTGGGAGACAGATTAAGCAAACTCCCTCGACATGATCCCTTTGA 2042  
Db 1 GCTCCTGGAGCTGGTGGGAGACAGATTAAGCAAACTCCCTCGACATGATCCCTTTGA 60

..

QY 2043 CCCAAGCTCTGCTCCTGCTCCTGACCAACCATGCCCTTTCTTAACTTCTCAAAACAGAT 2102  
Db 61 CCCAAGCTCTGCTCCTGCTCCTGACCAACCATGCCCTTTCTTAACTTCTCAAAACAGAT 120  
QY 2103 ACCAGGSCCTAAACTGCTTTTACCTCCCTCTACTAGTCAGGTAGTGTGGGAGGTC 2162  
Db 121 ACCAGGSCCTAAACTGCTTTTACCTCCCTCTACTAGTCAGGTAGTGTGGGAGGTC 180  
QY 2163 ACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAAGTCATGTGGGTATGTCG 2222  
Db 181 ACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACAAAGTCATGTGGGTATGTCG 240  
QY 2223 GGTAGAGAGAGGGGTAGCAAGTTTCTCTCTTGGTTCACATATCTCCCAAAGCTC 2282  
Db 241 GGTAGAGAGAGGGGTAGCAAGTTTCTCTCTTGGTTCACATATCTCCCAAAGCTC 300  
QY 2283 TGATCCCTGCCATGGGAAGTGGACAGAAACATGAGTCACTGACCTGCAGGCATCTTTAC 2342  
Db 301 YGATCCCTGCCATGGGAAGTGGACAGAAACATGAGTCACTGACCTGCAGGCATCTTTAC 360  
QY 2343 TGCAGCTCTGCCGCCCTGGAGGGGAGGGGAGGAAGTATGCGCTGCACATTTCT 2402  
Db 361 TGCAGCTCTGCCGCCCTGGAGGGGAGGGGAGGAAGTATGCGCTGCACATTTCT 420  
QY 2403 GAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGT 2462  
Db 421 GAGGCTACTGCATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGT 480  
QY 2463 TTGGGCCCGATAGGAAGTTCTCCGTGGCTCCCTCAGGAGAGGAGGAGGCTGAC 2522  
Db 481 TTGGGCCCGATAGGAAGTTCTCCGTGGCTCCCTCAGGAGAGGAGGAGGCTGAC 540  
QY 2523 ATTGCCAGTCTCTTCTGGGGCCCAAGCAGGTTGCAGGAGATCCAAATCCCATAGACAGCT 2582  
Db 541 ATTGCCAGTCTCTTCTGGGGCCCAAGCAGGTTGCAGGAGATCCAAATCCCATAGACAGCT 600  
QY 2583 C 2583  
Db 601 C 601

RESULT 7  
US-09-794-257-15  
; Sequence 15, Application US/09794257  
; Patent No. US2002009804A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1e1  
; FILE REFERENCE: Human G-Proteins  
; CURRENT APPLICATION NUMBER: US/09/794,257  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/185,606  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 576  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-794-257-15

Query Match 21.5%; Score 576; DB 10; Length 576;  
Best Local Similarity 100.0%; Pred. No. 1.7e-92;  
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 ATGCTTCTGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCAAATCAAGACGGG 197  
Db 1 ATGCTTCTGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCAAATCAAGACGGG 60  
QY 198 GCCTTCTGTGCGGAACCTTCATAGCCCGTGGGATAGACTTCAGGAACAAGGTGGT 257  
Db 61 GCCTTCTGTGCGGAACCTTCATAGCCCGTGGGATAGACTTCAGGAACAAGGTGGT 120

OY 258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGAGGAAACGGTTC 317  
DB 121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGAGGAAACGGTTC 180  
OY 318 CGAGCGTCACCCATGCTTTATACAGAGATGCTCAGGCGCTTGTCTCTGTATGACATC 377  
DB 181 CGAGCGTCACCCATGCTTTATACAGAGATGCTCAGGCGCTTGTCTCTGTATGACATC 240  
OY 378 ACCAAACAATCTCTTTCCACAACATCAGGCGCTGGCTCACTGAGATTCAATGATGCC 437  
DB 241 ACCAAACAATCTCTTTCCACAACATCAGGCGCTGGCTCACTGAGATTCAATGATGCC 300  
OY 438 CAGAGGAGCGTGGTATCATCTGCTAGGCAACRAGGCGGATATGAGCAGGCAAAAGAGTG 497  
DB 301 CAGAGGAGCGTGGTATCATCTGCTAGGCAACRAGGCGGATATGAGCAGGCAAAAGAGTG 360  
OY 498 ATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTACGGTGTTCCTTCTGTGAGACC 557  
DB 361 ATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTACGGTGTTCCTTCTGTGAGACC 420  
OY 558 AGCGCAAGACTGGCATCAATGTGGAGTTAGCTTTCTTGGCCATCGCCCAAGGAAGTGA 617  
DB 421 AGCGCAAGACTGGCATCAATGTGGAGTTAGCTTTCTTGGCCATCGCCCAAGGAAGTGA 480  
OY 618 TACCGGGCGGCGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677  
DB 481 TACCGGGCGGCGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540  
OY 678 CAGAAGAAGCGCTCAGCTGCTGCTTCATGTGA 713  
DB 541 CAGAAGAAGCGCTCAGCTGCTGCTTCATGTGA 576

RESULT 8

US-09-817-199A-25  
; Sequence 25, Application US/09817199A  
; Patent No. US20020142380A1  
; GENERAL INFORMATION:  
; APPLICANT: SHAO, Wei et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CLO01187  
; CURRENT APPLICATION NUMBER: US/09/817,199A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-817-199A-25

Query Match 21.4%; Score 573; DB 10; Length 601;  
Best Local Similarity 99.7%; Pred. No. 5.7e-92;  
Matches 573; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2055 GACCACCCATGCCCTTTCCTTTAACTTCTCAACAGATACACGAGGCGCTTAACTGCTTAC 2124  
DB 1 GACCACCCATGCCCTTTCCTTTAACTTCTCAACAGATACACGAGGCGCTTAACTGCTTAC 60  
OY 2125 CTCCTCCTACTAGTCAGGTAGGTGGTGGAGGTACCCATTTCCGAGTTAAACAA 2184  
DB 61 CTCCTCCTACTAGTCAGGTAGGTGGTGGAGGTACCCATTTCCGAGTTAAACAA 120  
OY 2185 TGCATATGAGTAAACAAAGTCAATGTGGGTATGCTGGGTAGAGAGGGGTAGCAAG 2244  
DB 121 TGCATATGAGTAAACAAAGTCAATGTGGGTATGCTGGGTAGAGAGGGGTAGCAAG 180  
OY 2245 TTCATGTCTCTCTTGGTGCATATCTCCAAAGCTCTGATCCCTGCGATGGGAAGTGG 2304  
DB 181 TTCATGTCTCTCTTGGTGCATATCTCCAAAGCTCTGATCCCTGCGATGGGAAGTGG 240

OY 2305 ACAGGAACATGAGCTCATGCTGAGGCATCTTACTGCAGCTCTGCGCGCCTGGAGG 2364  
DB 241 ACAGGAACATGAGCTCATGCTGAGGCATCTTACTGCAGCTCTGCGCGCCTGGAGG 300  
OY 2365 GGGAGAGGGGAGGAGAGATATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTCA 2424  
DB 301 RGGAGAGGGGAGGAGAGATATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTCA 360  
OY 2425 AGGCAAAATCTTCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCGATAAGAAAGTTCT 2484  
DB 361 AGGCAAAATCTTCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCGATAAGAAAGTTCT 420  
OY 2485 CCGTGGCCTCCCTCAGGCGAGCAGGAGGCTGACATTTGCCAGTCTCTTCTTGGGCC 2544  
DB 421 CCGTGGCCTCCCTCAGGCGAGCAGGAGGCTGACATTTGCCAGTCTCTTCTTGGGCC 480  
OY 2545 CAAGCAGGTTGACAGGATCCAAATCCCATAGACAGCTCTTGGGCCCTTTGCAATTTGAGTT 2604  
DB 481 CAAGCAGGTTGACAGGATCCAAATCCCATAGACAGCTCTTGGGCCCTTTGCAATTTGAGTT 540  
OY 2605 TTTCAGAAATTAACATGCACTATTTTGGAAAGCAA 2639  
DB 541 TTTCAGAAATTAACATGCACTATTTTGGAAAGCAA 575

RESULT 9

US-09-867-550-1811  
; Sequence 1811, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells a  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1811  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: (1)  
; OTHER INFORMATION: Wherein n is one of a or t or c or g  
US-09-867-550-1811

Query Match 16.6%; Score 443.4; DB 10; Length 447;  
Best Local Similarity 99.8%; Pred. No. 2.3e-69;  
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 GCGCGGCACTGCTCACCTCTCTGTCAGGAGACATGACGGGACCGCAGGCGCGTGGCCAC 70  
DB 3 GCGCGGCACTGCTCACCTCTCTGTCAGGAGACATGACGGGACCGCAGGCGCGTGGCCAC 62  
OY 71 CCGGATGCGAGGCGCGCGGAGCGCTCCCGCCCTGCGAGTCCGAGCTAGCAGCTCACGGG 130  
DB 63 CCGGATGCGAGGCGCGCGGAGCGCTCCCGCCCTGCGAGTCCGAGCTAGCAGCTCACGGG 122  
OY 131 CAAGTGTATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTCCCTGATCCAATCAA 190  
DB 123 CAAGTGTATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTCCCTGATCCAATCAA 182  
OY 191 AGACGGGCGCTTCTCTGTCGGGAACCTTCATAGCCACCGCTGGCATAGACTTCAGGAACAA 250

```
Db 183 AGAGGGGGCTTCTGTCCGGAACCTTCATAGCCACCGTGGCATAGACTTCAGGAACAA 242
QY 251 GGTGGTGACTGTGGATGGCTGAGATGAAGCTGAGATCTGGGACACCGCTGGGCGAGGA 310
Db 243 GGTGGTGACTGTGGATGGCTGAGATGAAGCTGAGATCTGGGACACCGCTGGGCGAGGA 302
QY 311 ACGTTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCCTTGCTGCTGTA 370
Db 303 ACGTTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCCTTGCTGCTGTA 362
QY 371 TGACATCACCACAACAAATCTTCTTCACACACATCAGGGGCTTGCTCAGTGAATTCATGA 430
Db 363 TGACATCACCACAACAAATCTTCTTCACACACATCAGGGGCTTGCTCAGTGAATTCATGA 422
QY 431 GTATGCCAGAGGAGCTGGTGATC 455
Db 423 GTATGCCAGAGGAGCTGGTGATC 447

RESULT 10
US-09-764-877-2275/c
; Sequence 2275, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2275
; LENGTH: 21833
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2275

Query Match 7.5%; Score 200.6; DB 10; Length 21833;
Best Local Similarity 74.1%; Pred. No. 9.9e-27;
Matches 254; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1525 TTTTGTGTTTTTTTAAATGAGTCTCGTTCGTGCGCCAGGCTCAGGTGCAGTAG 1584
Db 5032 TTTTGTGTTTTTTTAAATGAGTCTCGTTCGTGCGCCAGGCTCAGGTGCAGTAG 1584
QY 1585 TGAATCTCGGCTCACTACAACTCCACTCCCTGGGGCTCAAGCGATCCTCCACTCAG 1644
Db 4972 TGAATCTCGGCTCACTACAACTCCACTCCCTGGGGCTCAAGCGATCCTCCACTCAG 1644
QY 1645 CCGCCGAAGTAGTGGGACTATAGTGTGTACCATCACACCTGGGTAATTTTGTATTTT 1704
Db 4912 CCGCCGAAGTAGTGGGACTATAGTGTGTGTACCATCACACCTGGGTAATTTTGTATTTT 1704
QY 1705 TTGTAGACACAGGTTTCCCATGTTGCCAGGTGCTTGAATTCGTAGCTCAAGCA 1764
Db 4852 TAGTAGACACAGGTTTCCCATGTTGCCAGGTGCTTGAATTCGTAGCTCAAGCA 1764
QY 1765 ACTGCGCGGCTCGGCTCCCAAGTACTGGATTACAGCGAGAGGACCATGCCAGG 1824
Db 4792 ATCCACCTGCTGCGGCTCCCAAGTACTGGATTACAGCGAGAGGACCATGCCAGG 1824
QY 1825 CTAGATGTGCTTATCCCAATCCTTTGGCAGGCTGAGCTCC 1867
Db 4732 CCGTGAATTTATTCATGATATTTCTTTGATGAAAGGCCCTCC 4690

RESULT 11
US-09-764-877-3373/c
; Sequence 3373, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3373
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3373

Query Match 7.5%; Score 199.8; DB 10; Length 1130;
Best Local Similarity 77.1%; Pred. No. 9.2e-27;
Matches 243; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1516 GGAAAAGAGTTTTTTTAAATGAGTCTCGTTCGTGCGCCAGGCTGAG 1575
Db 673 GTAATTAATTTTCCCTTTTTTTTGGACAGAGGSCCTAGCTCTGACACCCAGGCTGAA 614
QY 1576 GTGCAGTAGTGAATCTCGGCTCACTACAACTCCACTCCCTGGGGCTCAAGGATCCTC 1635
Db 613 GTGCAGTGGTGTGATCTCAGCTCACTGCAACCTCTGCTTCCCTGGACTCAAGTGATCCTC 554
QY 1636 CCACCTCAGCGCCGGAAGTAGCTGGGACTATAGTGTGTACCATCACACCTCGCTAAATTT 1695
Db 553 CCACCTCAGCGCCCTTAAGTAGCTGGGACTTACAAATGCACACCCACACCTGGTTAATTT 494
QY 1696 TTGTATTTTGTAGACACAGGGTTTGGCCATGTTGCCAGGCTGGTCTTGAATTCCTGA 1755
Db 493 TTGTATTTTGTAGACACAGAGTTTGGCACGTTGGCCAGGCTGGTCTGGAACCTCTGG 434
QY 1756 GCTCAAGCAACCTCGCGGCTCGGCTCCCAAGTACTGGGATTTACACCCAGGACACC 1815
Db 433 GCTCAAGCGATCCGCCACCTTGGCTCCCAAGTCTAGGATTACAGGAGTGAGCCACA 374
QY 1816 ATGCCCGAGGCTAGAT 1830
Db 373 ATGCCCTGGCCGCAAT 359

RESULT 12
US-09-764-877-3236
; Sequence 3236, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3236
; LENGTH: 15745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3236

Query Match 7.4%; Score 198.6; DB 10; Length 15745;
Best Local Similarity 76.7%; Pred. No. 2.1e-26;
Matches 243; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1514 TAGGAAAAGAGTTTTTTTAAATGAGTCTCGTTCGTGCGCCAGGCTG 1573
Db 6486 TAAAGTTACTAACATTTTTTTTTTTTGGACAGGATCTTGTCTGTACCCAGGCTG 6545
QY 1574 AGGTGCAGTAGTGAATCTCGGCTCACTACACCTCCACTCCCTGGGGCTCAAGGATCC 1633
Db 6546 GAGTCAGTGGTACGATCTCAGCTCACTAGCCCTTAACCCACCAGGCTTATGCGCTCC 6605
```



Db	2518	AAATGGAATTCCTCTTTTTTTTTTTTGAAGAGGCTCATTTCTATCGCCCAAGCTGGAGT	2859
Qy	1578	GCAGTAGTGCAAATCTCCGCTCACTCAACCTCCACTCCCTGGGCTCAAGGAGTCTCTCC	1637
Db	2858	GCACTGGGTGATCCAGCTCACTGGAACCTCCGCTTCC -GGGGTCAGTGATCTCTCC	2800
Qy	1638	ACCTCAGCCCGCCGAAGTAGCTGGGACTATAGTGTTGTACCATCACACTGGCTAATTTTT	1697

